On Granularity in Information Systems Based on Binary Relation

Weihua Xu\textsuperscript{1}, Shihu Liu\textsuperscript{1}, Xiaoyan Zhang\textsuperscript{1}, Wenxiu Zhang\textsuperscript{2}
\textsuperscript{1}School of Mathematics and Statistics, Chongqing University of Technology, Chongqing, China
\textsuperscript{2}School of Science, Xi'an Jiaotong University, Xi'an, China
E-mail: chxuwh@gmail.com, liush02@126.com
Received January 21, 2011; revised March 15, 2011; accepted March 20, 2011

Abstract

In this paper, some important issues of granularity are discussed mainly in information systems (ISs) based on binary relation. Firstly, the vector representation method of knowledge granules is proposed in an information system based on binary relation to eliminate limitations of set representation method. Secondly, operators among knowledge granularity are introduced and some important properties of them are studied carefully. Thirdly, distance between two knowledge granules is established and granular space is constructed based on it. Fourthly, axiomatic definition of knowledge granularity is investigated, and one can find that some existed knowledge granularities are special cases under the definition. In addition, as an application of knowledge granular space, an example is employed to validate some results in our work.

Keywords: Binary Relation, Granular Space, Information System, Rough Set

1. Introduction

Rough set theory, proposed by Pawlak in the early 1980s [16], is an extension of the classical set theory and can be regarded as a soft computing tool to deal with uncertainty or imprecise information. It was well known that this theory is based upon the classification mechanism, in which case the classification can be viewed as an equivalence relation and knowledge granules induced by the equivalence relation can be viewed as a partition of universe. For this reason, it has been applied widely and successfully in feature selection [22], uncertainty reasoning [6], granular computing [9,14,29-33], data analysis [15,17,18] and data mining [24-27], etc.

Equivalence relation, as an important and primitive concept in Pawlak’s original rough set theory, still has many limitations. In order to eliminate these limitations resulted from equivalence relation and broaden its application fields, some meaningful works have been done in the past [3,4,10,21,28]. A knowledge granule, which is viewed as a partition of universe, plays an important role in investigating the information system. In ref. [33], L. A. Zadeh thought that nearly every field was permeated by granule and Hobbs discussed some properties with respect to knowledge granules in ref. [5]. In addition, L. A. Zadeh [34] pointed out that granulation is one of three basic concepts that underlie human cognition; the other two are organization and causation. Informally, granulation involves decomposition of whole into parts, organization involves integration of parts into whole and causation involves association of causes with effects. Hence, how to characterize the process of granulation has been a crucial problem. In other words, the validity of distinguishable ability, which is used to create the knowledge granules, should be examined because the knowledge granules in an information system are finite. Shannon [20], Beaubouef [1], Qian [19] and Liang [11,12] etc. used some useful methods to evaluate the uncertainty of information and L. A. Zadeh applied the notion of granularity to do this work, which presents a more visual and easily understandable description for a partition on the universe. Moreover, the relationships between several measures on knowledge in an information system were discussed in ref. [13]. These measures include granulation measure, information entropy, rough entropy, and knowledge granulation. Especially, closely associated with granularity, Xu [23] carefully discussed the properties of every granularity mentioned above. It is known that these measures have become effective mechanisms for evaluating uncertainty in rough set theory.

In this paper, our main contribution is to study the re-
lations among knowledge granules and the distinguish-
able ability of general binary relation which was used to
create the knowledge granules by the vector representa-
tion method in an information system. Firstly, four op-
erators among knowledge granules expressed as vectors
are proposed and the granular distance is defined. Then,
an axiomatic definition of knowledge granularity is pro-
bred by introducing a new binary relation.

The rest of this paper is organized as follows. Some
preliminary concepts required in our work are briefly
recalled in section 2. In section 3, four operators among
knowledge granules are introduced and important prop-
erties of them are acquired. While the focus of section 4
is the distance between any two knowledge granules and
granular space is constructed based on the distance.

Definition of knowledge granularity is proposed in sec-
tion 5 and some of its important properties are discussed.
The validity of some results obtained is examined by
introducing an example in section 6.

2. Preliminaries

In this section, we shall begin our work with some nec-


t essary concepts required in the sequel of this paper.
Detailed description of the theory can be found in refs.
[11,12,23].

Definition 2.1([37]) An information system is a tetra-

d \( I = (U, A, V, f) \), where

- \( U = \{u_1, u_2, \ldots, u_n\} \) is a non-empty finite set of ob-
jects called universe.
- \( A = \{a_1, a_2, \ldots, a_m\} \) is a non-empty finite set of at-
tributes.
- \( V = \bigcup_{a_i \in A} V_{a_i} \), \( V_{a_i} \) is the domain of attribute.
- \( f : U \times A \to V \), called an information function,
\( \forall a_i \in A, x \in U, f(x, a_i) \in V_{a_i} \).

An information system with decision is a special case of
an information system \( I = (U, A, V, f) \), in which
case attribute set \( A = C \cup D \) is the union of condi-
tional attributes \( C \) and decision attributes \( D \), with \( C \cap D = \phi \).

Any subset \( R \) of \( U \times U \) is called a relation on \( U \). For
any \( (x, y) \in U \times U \), if \( (x, y) \in R \), we say \( x \) has relation
\( R \) with \( y \), and denote this relation as \( xRy \). For an infor-
mation system \( I = (U, A, V, f) \) and \( B \subseteq A \), if denote
\( R_B = \{(x, y) | f(x, a) \neq f(y, a), x, y \in U, \forall a \in B \} \), then
\( R_B \) means a general binary relation with respect to \( B \)
on \( U \) and the system is a general information system where
“\( * \)” be “\( \geq \)”, “\( \leq \)” or “\( = \)”. Obviously, \( R_B \) is an equiva-
lence relation and the system is classical information
system when “\( * \)” be “\( = \)”.

Remark 1 Unless otherwise specified, information
systems appeared in the subsequent sections are general
information systems.

Let \( [u_j]_B = \{u_j | (u_j, u_j) \in R_B \} \) and \( U/R_B = \{[u_j]_B \}
\|u_j \in U\} \). \( U/R_B \) would be called a knowledge granules
of \( U \) with respect to attribute set \( B \).

Definition 2.2 ([35,38]) Let \( I = (U, A, V, f) \) be an
information system and \( B_1, B_2 \subseteq A \).

1) \( \forall u_i \in U \), if \( [u_i]_{B_1} = [u_i]_{B_2} \), we say that \( U/R_{B_1} \)
equal to \( U/R_{B_2} \), denoted by \( U/R_{B_1} = U/R_{B_2} \).
2) \( \forall u_i \in U \), if \( [u_i]_{B_1} \subseteq [u_i]_{B_2} \), we say that \( U/R_{B_1} \)
is finer than \( U/R_{B_2} \), denoted by \( U/R_{B_1} \leq U/R_{B_2} \).
3) If \( U/R_{B_1} \leq U/R_{B_2} \) and \( [u_i]_{B_1} \neq [u_i]_{B_2} \) for some
\( u_i \in U \), we say that \( U/R_{B_1} \) is properly finer than \( U/R_{B_2} \), denoted by \( U/R_{B_1} < U/R_{B_2} \).

Definition 2.3 ([35,38]) Let \( I = (U, A, V, f) \) be an
information system, \( X \subseteq U \) and \( B \subseteq A \). Given
\( R_B(X) = \{u | u \notin X, u \in U\} \),
\( \overline{R_B(X)} = \{u | u \notin X, u \in U\} \),
then \( R_B(X) \) and \( \overline{R_B(X)} \) are called the lower
and upper approximate sets of \( X \), respectively, with respect to \( B \).

Definition 2.4([2]) Let \( E \) be a non-empty set and “\( \leq \)”
be a binary relation on \( E \). If “\( \leq \)” satisfies the following
properties
1) \( \forall x \in E \), \( x \leq x \); \hspace{1cm} (Reflexive)
2) \( \forall x \in E \), if \( x \leq y \) and \( y \leq x \), then \( x = y \); \hspace{1cm} (Anti-symmetric)
3) \( \forall x, y, z \in E \), if \( x \leq y \) and \( y \leq z \), then \( x \leq z \); \hspace{1cm} (Transitive)
then the binary relation “\( \leq \)” is called a partial order and
the non-empty set \( E \) is a partially ordered set or a poset,
denoted by \( (E, \leq) \).

Definition 2.5 ([36]) Let \( (L, \leq) \) be a poset, if there
exist two operators \( \wedge \) and \( \vee \) on \( L : L^2 \to L \) such that
\( a \wedge b = b \wedge a \),
\( a \vee b = b \vee a \);
\( a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \),
\( a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \);
then \( L \) is called a lattice.
And if
\( a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c) \),
\( a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c) \);
then \( L \) is called an assignment lattice.
Furthermore, if
\( \forall a \in L, \exists a', s.t., (a') = a \) and \( a \leq b \Rightarrow b' \leq a' \),
Copyright © 2011 SciRes.
then \( L \) is called a complemented lattice.

**Definition 2.6** ([7, 8]) Let \( X \) be a non-empty finite set, and for any \( x, y \in X \), if there exists one real number \( d(x, y) \), such that

1) \( d(x, y) \geq 0 \), and \( d(x, y) = 0 \) iff \( x = y \);

\[ \text{(Non-negative)} \]

2) \( d(x, y) = d(y, x) \);

\[ \text{(Symmetry)} \]

3) \( d(x, y) \leq d(x, z) + d(z, y), z \in X \);

\[ \text{(Triangle inequality)} \]

then we have that \( X \) is a distance space, denoted by \((X, d)\), and \( d(x, y) \) represents the distance between \( x \) and \( y \).

**Example 2.1** Consider an ordered information system in Table 1.

From the table we can have a dominance relation \( R_A \) with respect to attribute set \( A = \{a_1, a_2, a_3, a_4, a_5, a_6\} \) and \( \{u_{ij}\} = \{u_{ij}(f(u_j, a_i) \geq f(u_j, a_i)(\forall a_i \in A)\} \), in which case we have that

\[
\begin{align*}
\{u_{ij}\}_A &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_B &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_C &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_D &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}.
\end{align*}
\]

If take \( B = \{a_1, a_2, a_3\} \), we have that

\[
\begin{align*}
\{u_{ij}\}_A &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_B &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_C &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_D &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}.
\end{align*}
\]

In addition, let \( C = \{a_2, a_4, a_5\} \), we have that

\[
\begin{align*}
\{u_{ij}\}_A &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_B &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_C &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}, \\
\{u_{ij}\}_D &= \{u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}, u_{ij}\}.
\end{align*}
\]

It is obviously that \( U/R_A < U/R_B \) and \( U/R_A < U/R_C \), which mean that knowledge granules \( U/R_A \) is finer than knowledge granules \( U/R_B \) and \( U/R_C \).

### 3. Operators among Knowledge Granules

From section 2, we can find that set representation method of knowledge granules doesn’t reflect all information included in the knowledge granules, because two different objects may have the same neighborhood. However, neighborhood of every different object makes important rules in the knowledge granule \( U/R_B \). In order to eliminate shortages of set representation method, we can use vector representation method to express the knowledge granules in this section.

**Definition 3.1** Let \( I = (U, A, V, f) \) be an information system based on binary relation, and \( B \subseteq A \). Vector \( \{[u_{i1}], [u_{i2}], \ldots, [u_{in}]\} \) is called vector representation of knowledge granule \( U/R_B \), denoted by \( K_B \). Obviously, dimension of knowledge granules \( K_B \) is the cardinality of universe \( U \) and it’s any correspondent component is just the neighborhood of each object in \( U \).

Generally, there may exist many knowledge granules with respect to a given information system. For the sake of simplicity, the notation “GS”, named “granular cluster”, is proposed to denote all the knowledge granules of information system \( I = (U, A, V, f) \). That is,

\[
\text{GS} = \left\{ K_B | B \subseteq 2^A, i = 1, 2, \ldots, 2^k \right\}.
\]

What is more, one knowledge granule can be denoted by \( K_B \) if and only if the general binary relation, denoted by \( R_B \), is an empty relation, i.e., \( K_B = (\emptyset, \emptyset, \ldots, \emptyset) \). Correspondingly, one knowledge granule can be denoted by \( K_B \) if and only if the general binary relation, denoted by \( R_B \), is a full relation, that is, \( K_B = (U, U, \ldots, U) \). And, the knowledge granule can be denoted by \( K_B \) and the corresponding general binary relation is \( R_B \).

**Definition 3.2** Let \( I = (U, A, V, f) \) be an information system based on binary relation, and \( K_B, K_B, K_B, K_B \in \text{GS} \).

1) If \( [u_{ij}]_{K_B} = [u_{ij}]_{K_B} \) for any \( u_i \in U \), then we say that \( K_B \) is isomorphic to \( K_B \), denoted by \( K_B = K_B \).

2) If \( [u_{ij}]_{K_B} \subseteq [u_{ij}]_{K_B} \) for any \( u_i \in U \), then we say that \( K_B \) is finer than \( K_B \), denoted by \( K_B \leq K_B \) or \( K_B \geq K_B \).

3) If \( K_B \geq K_B \), and \( [u_{ij}]_{K_B} \neq [u_{ij}]_{K_B} \) for some \( u_i \in U \), then we say that \( K_B \) is properly finer than \( K_B \), denoted by \( K_B \geq K_B \) or \( K_B > K_B \).

4) If \( [u_{ij}]_{K_B} \subseteq [u_{ij}]_{K_B} \) for some \( u_i \in U \) but \( [u_{ij}]_{K_B} \supseteq [u_{ij}]_{K_B} \) for other \( u_i \in U \), then we say that the relation between \( K_B \) and \( K_B \) is vague, denoted by \( K_B \wedge K_B \).

From above, we have that there exist four relations (equal, finer, properly finer and vague) among knowledge granules in the granular cluster GS.

---

**Table 1.** An ordered information system.

<table>
<thead>
<tr>
<th>( U )</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( a_3 )</th>
<th>( a_4 )</th>
<th>( a_5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u_1 )</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( u_2 )</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>( u_3 )</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>( u_4 )</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( u_5 )</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Theorem 3.1 \((G_S, \leq)\) is a poset.

Proof. 1) (Reflexive) For any \(K_B \in G_S\), that is \(K_B = \{[u_1], [u_2], \ldots, [u_n]\}\).

Since \([u_1] \leq [u_1] (\forall u_i \in U)\), we have that \(K_B \leq K_B\).

2) (Anti-symmetric) Let \(K_B, K_C \in G_S\). If \(K_B \leq K_C\) and \(K_C \leq K_B\), then \([u_1] \subseteq [u_1]_B\) and \([u_1]_C \subseteq [u_1]_C\) for any \(u_i \in U\). By Definition 3.2, so \([u_1]_B = [u_1]_C\), that is \(K_B = K_C\).

3) (Transitive) Let \(K_{B_1}, K_{B_2}, K_{B_3} \in G_S\). If \(K_{B_1} \leq K_{B_2}\) and \(K_{B_2} \leq K_{B_3}\), then \([u_1]_{B_1} \subseteq [u_1]_{B_2}\) and \([u_1]_{B_2} \subseteq [u_1]_{B_3}\) for any \(u_i \in U\). By Definition 3.2, we have that \([u_1]_{B_1} \subseteq [u_1]_{B_3}\), i.e., \(K_{B_1} \leq K_{B_3}\).

For a given information system, the knowledge granules can be induced by some relations, and they can also be obtained from known knowledge granules by operation. Hence, operator, as one of basic mathematical concepts, has to be mentioned. Operators in information systems can be divided into two types: operators among neighborhoods of objects and operators among knowledge granules. The former operators \((\cup, \cap, \sim)\) are based on classical sets while the later operators are performed through knowledge resolving or knowledge composing in essence. So, operators among knowledge granules should be proposed.

Definition 3.3 Let \(I = (U, A, V, f)\) be an information system based on binary relation, and \(K_B, K_C \in G_S\). For operators, denoted by \(\triangleleft, \supseteq, c\) and \(\ominus\), can be defined as follows.

1) \(K_B \cap K_C = ([u_1] \cap [u_1], [u_2] \cap [u_2], \ldots, [u_n] \cap [u_n])\).
2) \(K_B \cup K_C = ([u_1] \cup [u_1], [u_2] \cup [u_2], \ldots, [u_n] \cup [u_n])\).
3) \(K_B \oplus K_C = ([u_1] - [u_1], [u_2] - [u_2], \ldots, [u_n] - [u_n])\).
4) \(K_B^c = U - [u_1] = U - [u_1]\).

One knowledge granule can be generated from two or more different knowledge granules in granular cluster of an information system. For example, there exist another knowledge granule \(K_B \in G_S\) such that \(K_B = \cap K_B \cup K_B\) for \(K_B \in G_S\), then we say that the knowledge granule \(K_B\) is generated from \(K_B\) and \(K_B\). So as the instance of \(K_B = K_B \cap K_B\).

Proposition 3.1 Let \(I = (U, A, V, f)\) be an information system based on binary relation and \(K_B, K_B, K_B \in G_S\), then operators \(\triangleleft, \supseteq, c\) and \(\ominus\) satisfy the following properties.

1) \(K_B \cap K_B = K_B\).
2) \(K_B \cup K_B = K_B \cup K_B\). (Law of idempotency)
3) \(K_B \oplus K_B = K_B \oplus K_B\). (Law of commutation)
4) \(K_B \cap (K_B \cap K_B) = K_B\). (Law of association)
5) \(K_B \cup (K_B \cup K_B) = K_B\). (Law of absorption)
6) \(K_B^c = K_B^c\). (Law of double negative)
7) \((K_B \cap K_B)^c = K_B \cup K_B\) \(K_B \cup K_B\).
8) \(K_B \cup K_B^c = K_B\). (Law of zero or unity)

By Definition 3.2, one can find that the relation between every two knowledge granules can't always be characterized by finer or coarse, sometimes the relation may be vague in information systems. Therefore, we make a formal regulation for the relations among knowledge granules to give a more clear explanation.

Definition 3.4 Let \(I = (U, A, V, f)\) be an information system based on binary relation. \(K_B, K_B, K_B \in G_S\), \(K_B \leq \text{ finer}(K_B, K_B)\) means either \(K_B \leq K_B\) or \(K_B \leq K_B\). And \(K_B \geq \text{coarser}(K_B, K_B)\) means either \(K_B \geq K_B\) or \(K_B \geq K_B\).

From above, we can obtain the following results.

Proposition 3.2 Let \(I = (U, A, V, f)\) be an information system based on binary relation and \(K_B, K_B \in G_S\).

1) \(K_B \cap K_B \leq \text{ finer}(K_B, K_B)\).
2) \(K_B \cup K_B \geq \text{coarser}(K_B, K_B)\).

Proposition 3.3 Let \(I = (U, A, V, f)\) be an information system based on binary relation and \(K_B, K_B \in G_S\).

1) \(K_B \cap K_B = K_B\) and \(K_B \cap K_B = K_B\).
2) \(K_B \cup K_B = K_B\), and \(K_B \cup K_B = K_B\).
3) \(K_B \leq K_B\), then \(K_B \leq K_B\).
Theorem 3.2 Let \( I =(U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \). Then we can have that

1. \( K_B \cap K_C = K_B \) if and only if \( K_B \leq K_C \).
2. \( K_B \cup K_C = K_C \) if and only if \( K_B \leq K_A \).

Proof. It can be proved by Definition 3.2 and 3.3.

Theorem 3.3 Let \( I =(U, A, V, f) \) be an information system based on binary relation, then \((GS, \leq, \cap, \cup)\) is an assignment lattice.

Proof. By Theorem 3.1, we have that \((GS, \leq)\) is a poset. And terms (1), (2) and (4) in Definition 2.5 are obvious from (2), (3) and (5) in Proposition 3.1.

In addition, let \( K_B, K_C \in GS \), then for any \( u_i \in U \),

\[
K_B \cap K_C = K_C \iff [u_i]_B \cap [u_i]_C = [u_i]_B.
\]

\[
\iff [u_i]_B \subseteq [u_i]_C \iff K_B \leq K_C.
\]

\[
K_B \cup K_C = K_C \iff [u_i]_B \cup [u_i]_C = [u_i]_C.
\]

\[
\iff [u_i]_B \subseteq [u_i]_C \iff K_B \leq K_C.
\]

Thus, \((GS, \leq, \cap, \cup)\) is an assignment lattice.

Theorem 3.4 Let \( I =(U, A, V, f) \) be an information system based on binary relation, then \((GS, \leq, \cap, \cup, \wedge)\) is a complemented lattice.

Proof. By Theorem 3.3, we have that \((GS, \leq, \cap, \cup)\) is an assignment lattice. And from (6) in Proposition 3.1, one can get that \((K_B^c) = K_B^c\). Moreover, from (4) in Definition 3.3, one has that

\[
K_B \leq K_C \iff [u_i]_B \subseteq [u_i]_C \quad \text{for any } u_i \in U
\]

\[
\iff [u_i]_B \cap [u_i]_C = [u_i]_B
\]

\[
\iff K_B^c \leq K_C^c.
\]

The theorem was proved.

Example 3.1 (Continued from Example 2.1) From vector representation method we have that

\[
K_A = \{[u_1], [u_2], [u_3], [u_4], [u_5], [u_6]\},
\]

\[
K_B = \{[u_1, u_2], [u_3, u_4], [u_5, u_6]\},
\]

\[
K_C = \{[u_3], [u_4], [u_5, u_6]\}.
\]

Thus, by calculating, we obtain that

\[
K_B \cap K_C = \{[u_3], [u_5, u_6]\},
\]

\[
K_B \cup K_C = \{[u_1, u_2], [u_3, u_4], [u_5, u_6], [u_3, u_4], [u_5, u_6]\}.
\]

So, the following is obvious.

\[
K_B \cap K_C \subseteq \text{finer}(K_B, K_C),
\]

\[
K_B \cup K_C \supseteq \text{coarser}(K_B, K_C).
\]

Furthermore, by the Definition of “c” we have that

\[
K_B^c = \{[u_2, u_3, u_4], [u_1, u_5, u_6], [u_1, u_2, u_3, u_4], [u_1, u_2, u_3, u_4, u_5, u_6]\}.
\]

\[
K_C^c = \{[u_2, u_3, u_4], [u_1, u_5, u_6], [u_1, u_2, u_3, u_4], [u_1, u_2, u_3, u_4, u_5, u_6]\}.
\]

Obviously, \(K_B^c < K_C^c\).

4. Granular Space

Yao [30] proposed the concept of set closeness between two classical sets to measure the degree of the sameness of them. For the idea, distance between two different knowledge granules is investigated in this section to characterize the relationship among knowledge granules by the vector representation method. Moreover, we construct granular space with the distance to characterize the relationship among knowledge granules by the vector representation method.

Definition 4.1 Let \( I =(U, A, V, f) \) be an information system based on binary relation. Let \( \mathcal{H} \) be a map from GS to \([0, 1]^\|U\|\). For any \( K_B \in GS \), we define

\[
\mathcal{H}(K_B) = (h_{B_1}^p(u_1), h_{B_2}^p(u_2), \ldots, h_{B_n}^p(u_n)),
\]

where \( h_{B_i}^p(u_i) = \frac{|[u_i]_B^p|}{|V|} \). We call that the vector \( \mathcal{H}(K_B) \) is granular vector of knowledge granule \( K_B \), denoted by \( \overline{K_B} \), that is to say,

\[
\overline{K_B} = \mathcal{H}(K_B) = (h_{B_1}^p(u_1), h_{B_2}^p(u_2), \ldots, h_{B_n}^p(u_n)).
\]

Definition 4.2 Let \( I =(U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \). Four operators among granular vectors, denoted by \( \cap, \cup, \wedge, \vee \), can be defined as follows.

1. \( \overline{K_B} \cap \overline{K_C} = \overline{K_B \cap K_C} \)
2. \( \overline{K_B} \cup \overline{K_C} = \overline{K_B \cup K_C} \)
3. \( \overline{K_B} \wedge \overline{K_C} = \overline{K_B \cap K_C} \)
4. \( \overline{K_B} \vee \overline{K_C} = \overline{K_B \cup K_C} \)

In next, granular distance is proposed to measure relationship between any two knowledge granules.

Definition 4.3 Let \( I =(U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \)Granular distance between \( \overline{K_B} \) and \( \overline{K_C} \) is defined as

\[
d_p(\overline{K_B}, \overline{K_C}) = \frac{1}{\|\overline{K_B} - \overline{K_C}\|^p} \left( \sum_{u \in V} |h_{B_i}^p(u) - h_{C_i}^p(u)|^p \right)^{\frac{1}{p}}
\]
and \( p > 0 \).

So, \( d_p \) is Minkowski distance \([8]\).

In particular, if \( p = 1 \), then \( d_1 \) is Hamming distance which is

\[
d_1(K_B, K_C) = \frac{1}{|U|} \sum_{u \in U} |h_B^*(u) - h_C^*(u)|
\]

and if \( p = 2 \), then \( d_2 \) is Euclid distance which is

\[
d_2(K_B, K_C) = \frac{1}{|U|^{1/2}} \left( \sum_{u \in U} |h_B^*(u) - h_C^*(u)|^2 \right)^{1/2}
\]

Granular distance has the following important properties.

**Theorem 4.1** (Non-negativity) Let \( I = (U, A, V, f) \) be an information system based on binary relation. \( d_p(K_B, K_C) \geq 0 \) holds for any \( K_B, K_C \in GS \).

**Proof.** Let \( K_B, K_C \in GS \), we have that

\[
|h_B^*(u) - h_C^*(u)| \geq 0 \quad (u \in U)
\]

Then,

\[
d_p(K_B, K_C) = \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} |h_B^*(u) - h_C^*(u)|^p \right)^{1/p} \geq 0.
\]

The theorem was proved. \( \square \)

**Theorem 4.2** (Symmetry) Let \( I = (U, A, V, f) \) be an information system based on binary relation. \( d_p(K_B, K_C) = d_p(K_C, K_B) \) holds for any \( K_B, K_C \in GS \).

**Proof.** Let \( K_B, K_C \in GS \), we have

\[
d_p(K_B, K_C) = \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} |h_B^*(u) - h_C^*(u)|^p \right)^{1/p}
\]

\[
= \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} |h_C^*(u) - h_B^*(u)|^p \right)^{1/p}
\]

\[
d_p(K_C, K_B)
\]

The theorem was proved. \( \square \)

**Theorem 4.3** (Monotonicity) Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \).

1. If \( K_B \leq K_C \), \( K_B \neq K_B \), then
   \[
d_p(K_B, K_C) \leq d_p(K_C, K_B).
   \]

2. If \( K_B \leq K_C \), then \( d_p(K_B, K_C) \geq d_p(K_C, K_B) \).

**Proof.** (1) Since \( K_B \leq K_C \), we have that \( h_B^*(u) \leq h_C^*(u) \), where \( i = 1, 2, \cdots, |U| \). That is to say,

\[
0 \leq h_B^*(u) - \frac{1}{|U|} \leq h_C^*(u) - \frac{1}{|U|}.
\]

By \( K_B \neq K_C \), we can obtain that

\[
\left|h_B^*(u) - \frac{1}{|U|}\right| \leq \left|h_C^*(u) - \frac{1}{|U|}\right|
\]

Thus,

\[
d_p(K_B, K_C) = \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} \left|h_B^*(u) - \frac{1}{|U|}\right|^p \right)^{1/p}
\]

\[
\leq \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} \left|h_C^*(u) - \frac{1}{|U|}\right|^p \right)^{1/p}
\]

\[
d_p(K_C, K_B).
\]

2) It can be proved in the same way as (1). \( \square \)

**Theorem 4.4** (Invariability) Let \( I = (U, A, V, f) \) be an information system based on binary relation. \( d_p(K_B, K_C) = d_p(K_B, K_C) \) holds for any \( K_B, K_C \in GS \).

**Proof.** By Definition 3.3, 4.1 and 4.2, we have that

\[
\overline{K_B} = \left(1 - h_B^*(u_1), 1 - h_B^*(u_2), \cdots, 1 - h_B^*(u_{|U|})\right),
\]

\[
\overline{K_C} = \left(1 - h_C^*(u_1), 1 - h_C^*(u_2), \cdots, 1 - h_C^*(u_{|U|})\right).
\]

Therefore,

\[
d_p(K_B, K_C) = \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} \left|1 - h_B^*(u) - (1 - h_C^*(u))\right|^p \right)^{1/p}
\]

\[
= \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} \left|h_B^*(u) - h_C^*(u)\right|^p \right)^{1/p}
\]

\[
d_p(K_C, K_B).
\]

The theorem was proved. \( \square \)

**Theorem 4.5** (Boundedness) Let \( I = (U, A, V, f) \) be an information system based on binary relation. \( 0 \leq d_p(K_B, K_C) \leq 1 \) holds for any \( K_B, K_C \in GS \).

**Proof.** For any \( u \in U \), we have that

\[
0 \leq h_B^*(u) \leq 1 \quad \text{and} \quad 0 \leq h_C^*(u) \leq 1.
\]

That is, \( 0 \leq h_B^*(u) - h_C^*(u) \leq 1 \).

Therefore,

\[
0 \leq \frac{1}{|U|^{1/p}} \left( \sum_{u \in U} \left|h_B^*(u) - h_C^*(u)\right|^p \right)^{1/p} \leq 1.
\]

The theorem was proved. \( \square \)

In particular, one can obtain the following properties.

**Proposition 4.1** Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \).

\( d_p(K_B, K_C) = 0 \) if and only if \( K_B = K_C \).

**Proposition 4.2** Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \).

\( d_p(K_B, K_C) = 1 \) if and only if for some \( u \in U \), \( h_B^*(u) = 1 \) and \( h_C^*(u) = 0 \), while for other \( u \in U \), \( h_B^*(u) = 0 \) and \( h_C^*(u) = 1 \).

**Theorem 4.6** (Triangle inequality) Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C, K_B \in GS \), we have that
If there exist two nonnegative real numbers \( c_1 \) and \( c_2 \) which satisfy

1. Both of them are not equal to zero at the same time.
2. For any \( u_i \in U \),

\[
c_1 \left( b_{i2}(u_i) - b_{i1}(u_i) \right) = c_2 \left( b_{i1}(u_i) - b_{i2}(u_i) \right)
\]

Then, \( d_1(K_{i1}, K_{i2}) = d_2(K_{i1}, K_{i2}) + d_1(K_{i2}, K_{i1}) \).


Theorem 4.7 Let \( I = (U, A, V, f) \) be an information system based on binary relation. \((GS, d_p)\) is a distance space.

Proof. (1) The properties of Non-negativity and symmetry have been proved in Theorem 4.1 and 4.2.

(2) The property of triangle inequality has been proved in Theorem 4.6.

The theorem was proved. \( \square \)

Distance space \((GS, d_p)\) would be called generalized granular space, because every element in GS is knowledge granule and \( d_p \) is the distance between two knowledge granules.

Example 4.1 Let us consider the information system in Example 2.1.

By computing, we have that

\[
\begin{align*}
d_1(K_{i1}, K_{i2}) &= \frac{3}{25}, \\
d_1(K_{i2}, K_{i3}) &= \frac{3}{25}, \\
d_1(K_{i1}, K_{i3}) &= \frac{4}{25}, \\
d_2(K_{i1}, K_{i2}) &= \frac{1}{5}, \\
d_2(K_{i2}, K_{i3}) &= \frac{2}{5}, \\
d_2(K_{i1}, K_{i3}) &= \frac{3}{5}, \\
d_2(K_{i1}, K_{i2}) &= \frac{4}{5}.
\end{align*}
\]

when \( p = 1, 2 \) the following is obvious.

\[
\begin{align*}
d_p(K_{i1}, K_{i2}) &\leq d_p(K_{i1}, K_{i1}) + d_p(K_{i2}, K_{i1}), \\
d_p(K_{i1}, K_{i2}) &\leq d_p(K_{i1}, K_{i3}) + d_p(K_{i2}, K_{i3}), \\
d_p(K_{i1}, K_{i2}) &\leq d_p(K_{i1}, K_{i2}).
\end{align*}
\]

Furthermore, we can obtain that

\[
d_2(K_{i1}, K_{i2}) = \frac{115}{25} = d_2(K_{i2}, K_{i1}).
\]

If we take \( D = \{a_2, a_3, a_4, a_5\} \), then we can calculate

\[
\overline{K_D} = \left\{ \frac{2}{5}, \frac{2}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5} \right\}
\]

And the following holds

\[
K_1 \leq K_2 \leq K_3,
\]

\[
d_1(K_{i1}, K_{i2}) = d_1(K_{i2}, K_{i3}) + d_1(K_{i3}, K_{i2}).
\]

5. Knowledge Granularity

Intuitively, knowledge granules can represent the distinguishable ability of the general binary relation based on set of attributes in an information system. To some extent, the stronger its distinguishable ability is, the smaller the cardinality of every object’s neighborhood, while it is difficult to qualitatively depict distinguishable ability of some binary relations when relation among knowledge granules are vague. The reason is that the partial relation “\( \leq \)” only considers the inclusion relation of the neighborhood with the same sequence in knowledge granules. In order to eliminate the limitations and to discover the essence of distinguishable ability, a new binary
relation between knowledge granules, denoted by “\( \preceq \)”, is firstly introduced in this section.

In brief, \( K^c = \{ [u_{1j}]_{1}, [u_{2j}]_{2}, \ldots, [u_{nj}]_{n} \} \) is applied to denote one of new sequences of

\[
K_C = \{ [u_{1j}]_{1}, [u_{2j}]_{2}, \ldots, [u_{nj}]_{n} \},
\]

where \( n \) is the cardinality of \( U \).

**Definition 5.1** Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \). The new binary relation “\( \preceq \)” between knowledge granules is defined as follows.

1. If \( \|u_{ij}\|_B \leq \|u_{ij}\|_C \) for any \( j = 1, 2, \ldots, n \), then we say that \( K_B \) is finer than \( K_C \), denoted by \( K_B \preceq K_C \).
2. If \( \|u_{ij}\|_B \leq \|u_{ij}\|_C \) for any \( j = 1, 2, \ldots, n \) and \( \|u_{ij}\|_B \neq \|u_{ij}\|_C \) for some \( i \), then we say that \( K_B \) is properly finer than \( K_C \), denoted by \( K_B \preceq K_C \).
3. If \( \|u_{ij}\|_B = \|u_{ij}\|_C \) for any \( j = 1, 2, \ldots, n \), then we say that \( K_B \) is equal to \( K_C \), denoted by \( K_B \equiv K_C \).
4. If \( \|u_{ij}\|_B \neq \|u_{ij}\|_C \) for any \( j = 1, 2, \ldots, n \), then we say that \( K_B \) is identically unequal to \( K_C \), denoted by \( K_B \not\equiv K_C \).

**Theorem 5.1** Let \( I = (U, A, V, f) \) be an information system based on binary relation. We have that \( (GS, \preceq) \) is a poset.

**Proof.** The proof is similar to Theorem 3.1. □

From above, one can find that binary relation “\( \preceq \)” compares the relation of two knowledge granules with the cardinality of neighborhood. In a broad sense, it improved the limitations of the partial relation “\( \preceq \)”. By the depiction of “\( \preceq \)”, we have following properties.

**Corollary 5.1** The partial relation “\( \preceq \)” is a special instance of the relation “\( \preceq \)”.

**Corollary 5.2** Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \). We have that

1. \( K_B \cap K_C \preceq \) finer \( (K_B, K_C) \),
2. \( K_B \cup K_C \geq \) coarser \( (K_B, K_C) \)

**Definition 5.2** Let \( I = (U, A, V, f) \) be an information system based on binary relation. For any \( K_B \in GS \), if there exists a real number \( G_r(K_B) \) satisfying

1. \( G_r(K_B) \geq 0 \);
2. For any \( K_C \in GS \), \( G_r(K_B) = G_r(K_C) \) if \( K_B \equiv K_C \);
3. For any \( K_C, K_D \in GS \), \( G_r(K_B) \leq G_r(K_C) \leq G_r(K_D) \),
4. if \( K_B \preceq K_C \preceq K_D \), then \( G_r(K_B) \) is called a knowledge granularity of \( K_B \) with respect to \( B \) in the information system.

**Theorem 5.2** Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B, K_C \in GS \). \( K_B \approx K_C \) if and only if \( G_r(K_B) = G_r(K_C) \).

**Proof.** It can be proved directly by Definition 5.2. □

**Corollary 5.3** Let \( I = (U, A, V, f) \) be an information system based on binary relation, and \( K_B, K_C \in GS \). If \( K_B \approx K_C \), then \( G_r(K_B) = G_r(K_C) \).

**Remark 2** If \( G_r(K_B) = G_r(K_C) \), one couldn’t obtain \( K_B = K_C \) all the time.

**Example 5.1** Consider an information system \( I = (U, A, V, f) \) and \( U = \{u_1, u_2, u_3, u_4, u_5\} \). Take \( K_B = \{\{u_1\}, \{u_2, u_3\}, \{u_4\}, \{u_5\}\} \)

and \( K_C = \{\{u_1, u_2\}, \{u_3\}, \{u_4\}, \{u_5\}\} \).

Obviously, \( G_r(K_B) = G_r(K_C) \), but \( K_B \) isn’t equal to \( K_C \). In fact, we have that \( K_B \not\approx K_C \).

**Theorem 5.3** (Boundedness) Let \( I = (U, A, V, f) \) be an information system based on binary relation. \( G_r(K_B) \leq G_r(K_C) \leq G_r(K_A) \) holds for any \( K_B \in GS \).

**Proof.** From Definition 3.2 and Corollary 5.1, we have \( K_A \preceq K_B \preceq K_A \). Hence, \( G_r(K_A) \leq G_r(K_B) \leq G_r(K_A) \) holds by (3) of Definition 5.2.

The theorem was proved. □

From above, one can find that any knowledge granule has its upper bound and lower bound in an information system \( I = (U, A, V, f) \). In particular, if \( K_B, K_A \in GS \), then the following properties hold.

**Theorem 5.4** (Extremum) Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B \in GS \), the following results are true.

1. For any \( K_C \in GS \), \( G_r(K_B) = \max \{G_r(K_C)\} \) if and only if \( K_B = K_A \);
2. For any \( K_C \in GS \), \( G_r(K_B) = \min \{G_r(K_C)\} \) if and only if \( K_B = K_A \).

**Proof.** (1) Obviously, one can have for any \( K_C \in GS \), \( G_r(K_B) = \max \{G_r(K_C)\} \). If \( K_B = K_A \), then \( G_r(K_B) = \max \{G_r(K_C)\} \).

Otherwise, if \( G_r(K_B) = \max \{G_r(K_C)\} \) for any \( K_C \in GS \), it means \( G_r(K_C) \leq G_r(K_B) \). In particular, \( G_r(K_A) \leq G_r(K_B) \). We have \( K_B \preceq K_B \) by Definition 5.2. From Definition 5.1, \( K_B = K_A \) holds.

(2) The proof is similar to (1).

The theorem was proved. □

**Theorem 5.5** (Knowledge composed)

Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B \in GS \), where \( i = 1, 2, \ldots \). If a new knowledge granule, denoted by \( K_B \), can be com
posed of \( K_B \), i.e., \( K_B = \bigcup_i K_{B_i} \), then we have that \( G_i (K_B) \leq G_i (K_B) \).

**Proof.** Since \( K_B = \bigcup_i K_{B_i} \), we have that \( K_B \subseteq K_B \). So we can obtain \( G_i (K_B) \leq G_i (K_B) \) by Definition 5.2.

The theorem was proved. \( \square \)

**Theorem 5.6** (Knowledge resolved) Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B \in GS \). If knowledge granules \( K_B \) can be resolved into two or more new knowledge granules, denoted by \( K_B \), where \( i = 1, 2, \ldots \) that is to say, \( K_B = \bigcup_i K_{B_i} \), then we have \( G_i (K_B) \leq G_i (K_B) \).

**Proof.** The proof is similar to Theorem 5.5. \( \square \)

**Definition 5.3** (24) Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B \in GS \). Knowledge granulation of knowledge granule \( K_B \), which is denoted by \( G_{K_B} \), can be defined as

\[
G_{K_B} = \frac{1}{|U|} \sum_{u \in U} \left[ \left[ u \right]_{K_B} \right].
\]

From above, we can have that \( G_{K_B} \) still satisfies all the properties from Theorem 5.2 to Theorem 5.6.

**Corollary 5.3** Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( B \subseteq A \), then \( G_{K_B} \) in Definition 5.3 is a knowledge granularity under Definition 5.2.

**Definition 5.4** (24) Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B \in GS \). Rough entropy of knowledge granule \( K_B \), which is denoted by \( E_{K_B} \), can be defined as

\[
E_{K_B} = \left( \frac{1}{|U|} \right) \log \frac{1}{|\left[ u \right]_{K_B}|}.
\]

By definition of rough entropy, we can find that \( E_{K_B} \) still satisfies all the properties from Theorem 5.2 to Theorem 5.6.

**Corollary 5.4** Let \( I = (U, A, V, f) \) be an information system based on binary relation and \( K_B \in GS \), then \( E_{K_B} \) in Definition 5.4 is a knowledge granularity under Definition 5.2.

**Example 5.2** (Continued from Example 2.1)

By computing, we obtain that knowledge granularity after making operators \( \sqcup \) and \( \cap \) among different knowledge in the system of Example 2.1, which are expressed in Table 2 and 3, respectively.

\[
\text{Table 2. Operation on } \sqcup.
\]

<table>
<thead>
<tr>
<th>( \sqcup )</th>
<th>( K_d )</th>
<th>( K_B )</th>
<th>( K_C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_d )</td>
<td>0.24</td>
<td>0.36</td>
<td>0.40</td>
</tr>
<tr>
<td>( K_B )</td>
<td>0.36</td>
<td>0.36</td>
<td>0.52</td>
</tr>
<tr>
<td>( K_C )</td>
<td>0.40</td>
<td>0.52</td>
<td>0.40</td>
</tr>
</tbody>
</table>

By definition of knowledge granulation, we can have that \( G_{K_d} < G_{K_B} < G_{K_C} \).

Obviously, the following hold.

\[
G_{K_d} < G_{K_B} < G_{K_C} < G_{K_B \sqcup K_C}.
\]

In addition, construction of the Knowledge granularity can be illustrated from Figure 1.

**Remark 3** In this example we only consider the operation \( \sqcup \) and \( \cap \), respectively. Other operation, such as \( \ominus \) and \( \ast \), can be similarly considered.

**6. Case Study**

In order to illustrate the vector representation method of knowledge granules and knowledge granularity, the information system about CTR [38]{Car Test Results}, see Table 4, is introduced into this section.

The set of attributes \( A = \{a_1, a_2, \ldots, a_9\} \) in the system are showed as follows.

\[
\text{Table 3. Operation on } \cap.
\]

<table>
<thead>
<tr>
<th>( \cap )</th>
<th>( K_d )</th>
<th>( K_B )</th>
<th>( K_C )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_d )</td>
<td>0.24</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>( K_B )</td>
<td>0.24</td>
<td>0.36</td>
<td>0.24</td>
</tr>
<tr>
<td>( K_C )</td>
<td>0.24</td>
<td>0.24</td>
<td>0.40</td>
</tr>
</tbody>
</table>

\[
\text{Figure 1. Construction of knowledge granularity in Example 5.2.}
\]

\[
\text{Table 4. CTR (Car Test Results).
}\]

<table>
<thead>
<tr>
<th>U/A</th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( a_3 )</th>
<th>( a_4 )</th>
<th>( a_5 )</th>
<th>( a_6 )</th>
<th>( a_7 )</th>
<th>( a_8 )</th>
<th>( a_9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( u_1 )</td>
<td>c</td>
<td>6</td>
<td>y</td>
<td>E</td>
<td>m</td>
<td>h</td>
<td>h</td>
<td>a</td>
<td>m</td>
</tr>
<tr>
<td>( u_2 )</td>
<td>c</td>
<td>6</td>
<td>n</td>
<td>E</td>
<td>m</td>
<td>m</td>
<td>h</td>
<td>ma</td>
<td>m</td>
</tr>
<tr>
<td>( u_3 )</td>
<td>c</td>
<td>6</td>
<td>n</td>
<td>E</td>
<td>m</td>
<td>h</td>
<td>h</td>
<td>ma</td>
<td>m</td>
</tr>
<tr>
<td>( u_4 )</td>
<td>c</td>
<td>4</td>
<td>y</td>
<td>E</td>
<td>m</td>
<td>h</td>
<td>h</td>
<td>ma</td>
<td>1</td>
</tr>
<tr>
<td>( u_5 )</td>
<td>c</td>
<td>6</td>
<td>n</td>
<td>E</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>ma</td>
<td>m</td>
</tr>
<tr>
<td>( u_6 )</td>
<td>c</td>
<td>6</td>
<td>n</td>
<td>B</td>
<td>m</td>
<td>m</td>
<td>a</td>
<td>he</td>
<td></td>
</tr>
<tr>
<td>( u_7 )</td>
<td>c</td>
<td>6</td>
<td>n</td>
<td>E</td>
<td>m</td>
<td>m</td>
<td>h</td>
<td>ma</td>
<td>he</td>
</tr>
<tr>
<td>( u_8 )</td>
<td>s</td>
<td>4</td>
<td>n</td>
<td>B</td>
<td>sm</td>
<td>h</td>
<td>lo</td>
<td>ma</td>
<td>1</td>
</tr>
</tbody>
</table>
$a_1$ — size, over all length, $a_2$ — number of cylinders, 
$a_3$ — presence of a turbocharger, 
$a_4$ — type of fuel system, 
$a_5$ — engine displacement, $a_6$ — compression, 
$a_7$ — power, $a_8$ — type of transmission, 
$a_9$ — weight. 
And values of attributes mean as follows.

c — compact,    s — subcompact,  sm — small, 
n — no,        E — EFI,         B — 2-BBL, 
y — yes,        m — medium,    ma — manual, 
h — high,       he — heavy,      l — light, 
lo — low ,        a — auto.

Let and it will obtain an equivalence relation 
$[u_i]_{K_B} = \{u_j | f_i(u_j) = f_i(u_i) (\forall x_i \in B)\}.$ 
Thus, we have that 
$K_B = \{\{u_1\}, \{u_2,u_3,u_4,u_5\}, \{u_2,u_3,u_4,u_5,u_6\}, \{u_4\}, \{u_2,u_3,u_4,u_5,u_7\}\} 
$ and 
$\overline{K_B} = \left(\begin{array}{c}
1 4 4 1 4 1 4 1 \\
5 8 8 8 8 8 8 8
\end{array}\right).$

If we denote $C = \{a_1, a_2, a_3, a_4, a_6\}$, then we have that 
$K_C = \{\{u_1\}, \{u_2\}, \{u_3,u_4,u_5\}, \{u_2,u_3,u_4,u_5,u_6\}, \{u_4\}, \{u_2,u_3,u_4,u_5,u_7\}\} 
$ and 
$\overline{K_C} = \left(\begin{array}{c}
1 1 1 1 1 1 1 1 \\
5 5 5 5 5 5 5 5
\end{array}\right).$

Moreover, if we denote $D = \{a_1, a_2, a_3, a_6\}$, then we have that 
$K_D = \{\{u_1,u_2\}, \{u_2,u_3,u_4,u_5\}, \{u_2,u_3,u_4,u_5,u_6\}, \{u_2,u_3,u_4,u_5,u_7\}, \{u_3,u_4\}, \{u_2,u_3,u_4,u_5,u_6,u_7\}\}, 
$ and 
$\overline{K_D} = \left(\begin{array}{c}
2 3 1 2 3 1 3 1 \\
6 8 8 8 8 8 8 8
\end{array}\right).$

By computing, we can obtain the following properties. 
(1) $K_C < K_B$, $K_C < K_D$, $K_B \geq K_D$, $K_B \neq K_D$. 
(2) $K_C \sqcap K_D = K_C$, $K_C \sqcap K_B = K_C$. 
(3) $K_B \sqcap K_D \leq \text{finer } (K_B, K_D).$ 
(4) $K_B \sqcup K_D \geq \text{coarser } (K_B, K_D).$

And we have that 
$d_1(K_B, K_C) = 0.1875$,  $d_1(K_B, K_D) = 0.1250$, 
$d_1(K_C, K_B) = 0.1250$,  $d_2(K_B, K_C) = 0.2652$, 
$d_2(K_B, K_D) = 0.1654$,  $d_2(K_C, K_B) = 0.1654$

and 
\[
\overline{K_B} = \left(\begin{array}{c}
7 4 4 7 4 7 4 7 \\
8 8 8 8 8 8 8 8
\end{array}\right), \\
\overline{K_C} = \left(\begin{array}{c}
6 5 7 6 5 7 5 7 \\
8 8 8 8 8 8 8 8
\end{array}\right). 
\]

Obviously, the following hold. 
(5) $d_i(K_B, K_D) = 0.1250 = d_i(\overline{K_B}, \overline{K_D})$. 
(6) $d_i(K_B, K_C) \leq d_i(K_B, K_D) + d_i(K_C, K_D)$, 
$d_i(K_B, K_C) \leq d_i(K_B, K_C) + d_i(K_D, K_C)$, 
$d_i(K_C, K_D) \leq d_i(K_B, K_C) + d_i(K_B, K_D)$. 

where $i = 1, 2$.

Moreover, we can have Table 5 about knowledge granularity of the system in Table 4 by computing.

In what follows Figure 2 is received to illustrate construction of knowledge granules mentioned in system about CTR.

7. Conclusions

Rough set theory is a powerful soft computing tool to deal with uncertainty and imprecision information. How to represent knowledge granules in information systems based on binary relation is one of the important research

<table>
<thead>
<tr>
<th>$G_C$</th>
<th>0.3125</th>
<th>0.1250</th>
<th>0.2500</th>
<th>0.2188</th>
<th>0.3438</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_B$</td>
<td>$K_B$</td>
<td>$K_B$</td>
<td>$K_B$</td>
<td>$K_B$</td>
<td>$K_B$</td>
</tr>
<tr>
<td>$K_D$</td>
<td>$K_D$</td>
<td>$K_D$</td>
<td>$K_D$</td>
<td>$K_D$</td>
<td>$K_D$</td>
</tr>
<tr>
<td>$K_C$</td>
<td>$K_C$</td>
<td>$K_C$</td>
<td>$K_C$</td>
<td>$K_C$</td>
<td>$K_C$</td>
</tr>
<tr>
<td>$K_B$ $\sqcap$ $K_C$</td>
<td>$K_B$ $\sqcap$ $K_D$</td>
<td>$K_B$ $\sqcup$ $K_D$</td>
<td>$K_B$ $\sqcup$ $K_C$</td>
<td>$K_B$ $\sqcup$ $K_C$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Construction of knowledge granularity in Example CTR.
tasks. In this paper, the vector representation method is proposed to eliminate limitations of set representation method, in which case the granular space is constructed by defining the distance between any two knowledge granules. In addition, knowledge granularity is investigated and some of its important properties are discussed carefully. As an application of knowledge granular space, an example is applied to illustrate the validity of some results obtained in our work.

8. References


