Knowledge-Based Systems 22 (2009) 38–45

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Real formal concept analysis based on grey-rough set theory ☆

Qiang Wu [a,b,*], Zongtian Liu [b]

[a] Department of Computer Technology and Science, Shaoxing College of Art and Science, Shaoxing, Zhejiang 312000, PR China
[b] School of Computer Engineering and Science, Shanghai University, Shanghai 200072, PR China

ARTICLE INFO

Article history:
Received 15 February 2008
Received in revised form 10 May 2008
Accepted 10 June 2008
Available online 18 June 2008

Keywords:
Grey information sysytem
Grey-rough set
Formal concept analysis
Galois lattices construction

ABSTRACT

One of the main concepts in grey system theory is how systems should be controlled under incomplete or lack of information situation. Grey number denoting an uncertain value is described in real interval from this concept. In this paper, we introduce the real formal concept analysis based on grey-rough set theory by using grey numbers, instead of binary values. We propose, to extend the notion of Galois connection in a real binary relation as well as the notions of formal concept and Galois lattice. The relationships between the new notions and old ones are discussed. Finally, we present a grey-rough set approach to Galois lattices reduction.

1. Introduction

Grey system theory [1–4], proposed by Deng, covers grey classification [5–7], grey control, grey decision-making [8], grey prediction [9–13], grey structural modelling [14,15], grey relational analysis as well as grey-rough sets [8,16–18], etc. It deals with the uncertainty over how systems with incomplete or lack of information should be controlled. One of the important concepts is a grey number. It is a number whose exact value is unknown but range is known. One of the practical applications of grey numbers is in error analysis [19] using a form $x = x_{best} + \delta_x$, where x is the measured value, x_{best} is the best estimate and δ_x is the error in the measurement of x. This form is equal to $\otimes x \rightleftarrows [x_{best} - \delta_x, x_{best} + \delta_x],\ \tilde{\otimes} x = x_{best}$ in grey system theory: $\otimes x$ is the interval where the best estimate exists and $\tilde{\otimes} x$ is one of the exact values that seems to be the best estimate. In real applications, a solution of equations or an optimized parameter is $\tilde{\otimes} x$ and their condition given in advance is $\otimes x$. Thus grey system theory deals with uncertainty unlike those of fuzzy set theory or rough set theory. The grey lattice operation [5,7] is one of the operations for grey numbers that modifies a range of given intervals of grey numbers. It is more suitable to handling information tables containing interval data. With the motivation, Yamaguchi et al. [20] proposed a new rough set model named grey-rough set, which is a new collaboration of rough set theory and grey system theory. A grey-rough approximation is based on the grey lattice relation instead of an equivalence class and an indiscernibility relation in Pawlak's model. Compared with the classical rough set, the proposal extends a treatable value into interval data. It provides a maximum solution and minimum solution both in upper and lower approximations. It give us a new mathematical background to develop a data set containing interval data.

Wille, in his Formal Concept Analysis [21], proposes a theory that allows us to formalize the three basic ideas of the conceptual knowledge, the objects, the attributes and the concepts. These ideas are linked through three basic relations: one object has an attribute, one object belongs to a concept and one concept is a subconcept of another one. With this analysis, a model to represent the concepts and to set up hierarchies among them is defined by Wille. We notice that most existing work focuses only on binary data. In order to generalize this work, the Galois lattice formalism was extended to symbolic data by [22] and further developed by [23–25]. Nevertheless, the general formalism of Galois lattice was addressed by [26,27]. The rationale for this generalization is that nowadays, either descriptions of data are complex, or the size of datasets is drastically growing up so that if, for example, we want to deal with classes of data, we need descriptions that are much more complex than 0 or 1.

Polaillon and Diday [28] proposed an extension of two classical algorithms (Ganter and Chein) and an incremental one (Godin et al.) to multivariate, interval and histogram data with missing values. In [29], Baklouti et al. proposed a fast Galois lattice-building algorithm based on dichotomic search and working for objects having general description. Although the algorithm can deal with general data, they have to redefine the set, the order relation, the operation of infimum and the largest element in different processes. Jaoua and Elloumi [30] introduced the notion of a real set as an extension of a crisp and a fuzzy set by using sequences of intervals as membership degrees, instead of a single value in [0,1]. They also proposed, to extend the notion of Galois connection in a real binary relation as well as the notions of rectangular rela-

☆ This work is partly supported by the National Natural Science Foundation of China with Grant No. 60575035.
* Corresponding author.
E-mail address: cswq@zscas.edu.cn (Q. Wu).

0950-7051/$ - see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.knosys.2008.06.001

tion, formal concept and Galois lattice. They had applied the strict Galois connection to build a real classifier system. But, their Galios lattices are different from classical concept lattices. Its intent and extent of bottom nodes are both empty. The object and attribute universe are intent, extent of top node, respectively.

In this paper, a new formal concept analysis approach for interval data based on grey-rough set theory is proposed. Since it is suitable for interval data reduction of attributes, we also propose, to use grey-rough set approach reducing real Galois lattices.

2. Grey lattice operation and grey-rough sets

Let G be the universal set, x be an element of G ($x \in G$), R be the set of real numbers and $X \subseteq R$ be the set of value range that x may hold.

Definition 2.1 [20]. Let Gr be a grey set of G defined by two mappings of the upper membership function $\bar{\mu}_{Gr}(x)$ and the lower membership function $\underline{\mu}_{Gr}(x)$ as follows:

$$\bar{\mu}_{Gr}(x) : G \to [0,1]; \quad \underline{\mu}_{Gr}(x) : G \to [0,1]$$

where $\underline{\mu}_{Gr}(x) \leqslant \bar{\mu}_{Gr}(x)$ and $x \in G$. When $\underline{\mu}_{Gr}(x) = \bar{\mu}_{Gr}(x)$, the grey set becomes a fuzzy set, which means that grey system theory deals with fuzzy situation more flexibly.

Definition 2.2 [20]. When two values $\underline{x}$, $\bar{x}$ ($\underline{x} = \inf X, \bar{x} = \sup X$) are given in x, then x is defined using a form $\otimes x = x|_{\underline{x}}^{\bar{x}}$ as follows:

(1) If and only if $x \to -\infty$ and $x \to +\infty$, $\otimes x$ is called a black number.
(2) If and only if $\underline{x} = \bar{x}$, $\otimes x$ is called a white number or a whitened value, which is denoted by $\tilde{\otimes} x$.
(3) Otherwise, $\otimes x \rightleftarrows [\underline{x}, \bar{x}]$ is called a grey number.

In other words, interval data are grey numbers and non-interval data are white numbers. If assume that $\otimes x \rightleftarrows [\underline{x}, \bar{x}]$, $\otimes y \rightleftarrows [\underline{y}, \bar{y}]$, then we have following definitions.

Definition 2.3 [20]. Let '$\rightleftarrows$' denote equality for two grey numbers $\otimes x$ and $\otimes y$ called the grey lattice coincidence relation as follows: $\otimes x \rightleftarrows \otimes y$ if and only if $\underline{x} = \underline{y}$ and $\bar{x} = \bar{y}$.

This relation indicates that the two endpoints are equal at the same time, which is distinguished from '='.

Definition 2.4 [20]. Let '$\to$' denote inclusion for two grey numbers $\otimes x$ and $\otimes y$ called the grey lattice inclusion relation as follows: $\otimes x \to \otimes y$ if $\underline{y} \leqslant \underline{x}$ and $\bar{x} \leqslant \bar{y}$.

In grey system theory, the grey arithmetic operation and the grey lattice operation [2,8,17] are introduced for grey numbers. The operators *Join*($\vee$), *Meet*($\wedge$), *Complement*($\otimes x^c$) and *Exclusive Join*($\oplus$) are given for three grey numbers $\otimes x \rightleftarrows [\underline{x}, \bar{x}]$, $\otimes y \rightleftarrows [\underline{y}, \bar{y}]$ and $\otimes z \rightleftarrows [\underline{z}, \bar{z}]$ as follows:

$$\otimes x \vee \otimes y \rightleftarrows [\min(\underline{x}, \underline{y}), \max(\bar{x}, \bar{y})]$$

$$\tilde{\otimes} x \vee \tilde{\otimes} y \rightleftarrows [\min(\tilde{\otimes} x, \tilde{\otimes} y), \max(\tilde{\otimes} x, \tilde{\otimes} y)]$$

$$\otimes x \wedge \otimes y \rightleftarrows \begin{cases} [\underline{x}, \bar{x}] & \text{if } \otimes x \to \otimes y \\ [\underline{y}, \bar{y}] & \text{if } \otimes y \to \otimes x \\ [\underline{x}, \bar{y}] & \text{if } \underline{x} \to \otimes y \text{ and } \bar{y} \to \otimes x \\ [\underline{y}, \bar{x}] & \text{if } \underline{y} \to \otimes x \text{and } \bar{x} \to \otimes y \\ \varnothing & \text{otherwise} \end{cases}$$

$$\tilde{\otimes} x \wedge \hat{\otimes} y \rightleftarrows \begin{cases} \tilde{\otimes} x & \text{if } \tilde{\otimes} x = \tilde{\otimes} y \\ \varnothing & \text{otherwise} \end{cases}$$

$$\otimes x^c = \{x \in X^c | x < \underline{x}, \bar{x} < x\} X^c = G \setminus X$$

$$\otimes x \oplus \otimes y \rightleftarrows \begin{cases} (\otimes x^c \wedge \otimes y^c)^c & \text{if } \otimes x \wedge \otimes y \rightleftarrows \varnothing \\ (\otimes x \vee \otimes y) \wedge (\otimes x \wedge \otimes y)^c & \text{if } \otimes x \wedge \otimes y \not\rightleftarrows \varnothing \end{cases}$$

Obviously, the operators *Join*($\vee$) and *Meet*($\wedge$) satisfy associative law: $\otimes x \vee (\otimes y \vee \otimes z) \rightleftarrows (\otimes x \vee \otimes y) \vee \otimes z, \otimes x \wedge (\otimes y \wedge \otimes z) \rightleftarrows (\otimes x \wedge \otimes y) \wedge \otimes z$ and $\otimes x \wedge \otimes y \to \otimes x$.

Whitening functions [7,17] compute a whitened value from a grey number. Especially the meet operation, the diameter and the overlap are mainly used to make a discernibility matrix of grey-rough reduction.

Diameter $dia(\otimes x) = \bar{x} - \underline{x}$

Overlap $ov(\otimes x, \otimes y) = \dfrac{dia(\otimes x \wedge \otimes y)}{dia(\otimes x \vee \otimes y)}$

where

$$\otimes x \wedge \otimes y \rightleftarrows \varnothing \iff ov(\otimes x, \otimes y) = 0; \quad \otimes x \rightleftarrows \otimes y \iff ov(\otimes x, \otimes y) = 1$$

Let $IS = (G, M, V, f_{\otimes})$ denote an information system called a grey information system [8], where

- G: a set of objects called the universe
- M: a set of attributes
- V: a set of values, $V \subseteq R$ in this paper
- $f_{\otimes}$: the information function as $f_{\otimes}$: $G \times M - \to V$

A grey-rough approximation [7] for a grey information system IS is based on the meet operation and the grey lattice inclusion ($\to$). Let x be an object of G, a be an attribute of M and $f_{\otimes}(x,a) \rightleftarrows [\underline{f_{\otimes}}(x,a), \overline{f_{\otimes}}(x,a)]$ be a value which x holds on the attribute a, where an ordered pair $(x,a) \in G \times M$, $\underline{f_{\otimes}}(x,a) = \inf V_a$ and $\overline{f_{\otimes}}(x,a) = \sup V_a$. Let $f_{\otimes}(s,a) \rightleftarrows [\underline{f_{\otimes}}(s,a), \overline{f_{\otimes}}(s,a)]$ be a value on $a$called an objective of approximation; the upper approximation $GL^*(f_{\otimes}(s,a))$ and lower approximation $GL^*(f_{\otimes}(s,a))$ are given as follows:

$$GL^*(f_{\otimes}(s,a)) = \{x \in G | f_{\otimes}(x,a) \wedge f_{\otimes}(s,a) \not\rightleftarrows \varnothing\}$$
$$GL_*(f_{\otimes}(s,a)) = \{x \in G | f_{\otimes}(x,a) \to f_{\otimes}(s,a)\}$$

$GL(f_{\otimes}(s,a))$ is a single-attribute approximation on an attribute a of M.

A multi-attribute approximation is also given. Let $M = \{a_1, a_2, \ldots, a_n\}$ be a set of n attributes, $S = \{f_{\otimes}(s,a_1) f_{\otimes}(s,a_2), \ldots, f_{\otimes}(s,a_n)\}$ be a set of n values on attributes of M denoting an objective. The upper approximation $GW^*(S)$ and lower approximation $GW^*(S)$ are given as follows:

$$GW^*(S) \rightleftarrows [\underline{GW}^*(S), \overline{GW}^*(S)]$$
$$GW_*(S) \rightleftarrows [\underline{GW}_*(S), \overline{GW}_*(S)]$$

where

$$\underline{GW}^*(S) = \cap_{k=1}^{n} GL^*(f_{\otimes}(s, a_k))$$
$$\overline{GW}^*(S) = \cup_{k=1}^{n} GL^*(f_{\otimes}(s, a_k))$$
$$\underline{GW}_*(S) = \cap_{k=1}^{n} GL_*(f_{\otimes}(s, a_k))$$
$$\overline{GW}_*(S) = \cup_{k=1}^{n} GL_*(f_{\otimes}(s, a_k))$$

A pair of interval sets $\langle GW^*(S), GW^*(S)\rangle$ is a multi-attribute grey-rough set. The multi-attribute approximation is mainly used in approximation [7]. The single-attribute approximation is mainly used for reduction.

Example 1. Table 1 is a sample grey information. Each object $\{s, x_1, x_2, x_3, x_4\} = G$ including the objective holds three attributes $\{a_1, a_2, a_3\} = M$ as shown in Table 1, and then we have