Knowledge-Based Systems 135 (2017) 18–28

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A unified information measure for general binary relations

Changzhong Wang[a,*], Qiang He[b], Mingwen Shao[c], Yangyang Xu[a], Qinghua Hu[d]

[a] Department of Mathematics, Bohai University, Jinzhou, 121000, PR China
[b] College of Science, Beijing University of Civil Engineering and Architecture, Beijing, 100044, PR China
[c] College of Computer and Communication Engineering, Chinese University of Petroleum, Qingdao, Shandong 266580, PR China
[d] School of Computer Science and Technology, Tianjin University, Tianjin 300072, PR China

ARTICLE INFO

Article history:
Received 26 December 2016
Revised 7 June 2017
Accepted 15 July 2017
Available online 6 September 2017

Key words:
Binary relation
Rough set
Entropy
Uncertainty
Information system

ABSTRACT

Shannon's entropy and its variants have been applied to measure uncertainty in a variety of special binary relations. However, few studies have been conducted on uncertainty of general binary relations. In this study, we present a unified form of uncertainty measures for general binary relations. We redefine the concepts of entropy, joint entropy, conditional entropy, and mutual information. These uncertainty measures are generalizations of corresponding measures of special relations. We study the relationship between these measures and examine important properties. Finally, numerical experiments are performed to identify applications of the proposed uncertainty measures. Comparing with existing uncertainty measures, the proposed method not only addresses the uncertainty of heterogeneous data sets, but also exhibit better performance in attribute reduction. This study can provide a fundamental framework for uncertainty theories of special rough set models.

1. Introduction

Rough set theory has become a popular approach for dealing with vague and uncertain information. One major advantage of the theory is that it uses only internal knowledge and does not rely on prior model assumptions. In recent years, this theory has received wider attention as a means of data analysis and has been extensively applied in the fields of reasoning with uncertainty, attribute reduction, rule extraction, and classification learning [2–10,13–17,28–34,41–52].

Relations on a universe of discourse are the theoretical basis of rough set theory. Different rough set models employ different types of relations such as equivalence relations, similarity relation, dominance relation, and so on. Two core concepts in rough set models are granularity and approximation. Rough set models use a type of relation to group objects into different granules, which are called elementary information granules, and approximately describe a decision. Such approximation is always vague because of the uncertainty caused by granularity of representation of the decision information.

Information entropy, proposed by Shannon, has been an effective and powerful mechanism for characterizing uncertainty information. The extension of entropy and its variants has been applied in the rough set field [4–12,18–27,35–39,44,52]. For example, Duntch and Gediga introduced information entropy into rough sets and proposed three types of conditional entropies for predicting a decision attribute [9]. Beaubouuef et al. developed a method for measuring uncertainty of rough approximation and rough relation from data by introducing the concept of rough entropy [1]. Wierman presented the uncertainty measures of granularity by defining rough set entropy [36]. Liang et al. proposed a new method for evaluating both uncertainty and fuzziness of knowledge from an information system [18,19]. Hu redefined joint and conditional entropy based on Yager's entropy in a neighborhood approximation space to measure the uncertainty of rough approximation [12]. Slezak used information entropy to compute the uncertainty of equivalence relations and identify the relative reducts of a decision table [26]. Qian and Liang presented the concepts of combination entropy and combination granulation for evaluating uncertainty of a discrete variable in terms of the partition ability of knowledge [24]. Yao defined a granulation measure from the viewpoint of granulation [39]. Dai et al. proposed a new type of conditional entropy based on tolerance relation for an incomplete decision system [6–8]. These studies all examine the entropy of special binary relations. However, they consider only the single structural information of data while ignoring the structural information of heterogeneous data sets.

In fact, information systems based on general binary relations exist in many fields. For example, complex network systems, social evaluation surveys, and fragmentary ordered information systems are good examples of these types of information systems. Alterna-

* Corresponding author.
E-mail address: changzhongwang@126.com (C. Wang).

http://dx.doi.org/10.1016/j.knosys.2017.07.017
0950-7051/

tively, a data set may contain nominal, ordinal, and scale features. If nominal features induce equivalence relations, ordinal features induce ordered relations, and scale features induce similarity relations. We can then derive a complex heterogeneous information system, which can be considered to be based on general binary relations. This is because all existing theories about information entropy do not address mixed-type relations but consider only special types of binary relations. Furthermore, similarity, ordered, and other types of relations are only the special cases of general binary relations. Therefore, a study of uncertainty measures for general binary relations is crucial. It not only handles the uncertainty information of complex heterogeneous data sets, but can also unify the uncertainty theories of special rough set models into a theoretical framework. In this study, we address the issue of uncertainty measures for general binary relations and present a unified form of uncertainty measures: entropy and its variants for general binary relations. We redefine the concepts of entropy, joint entropy, conditional entropy, and mutual information. These uncertainty measures are generalizations of corresponding measures of special relations. We study the relationship between these measures and identify important properties. Finally, we conduct some experiments to identify applications of the proposed uncertainty measures.

The remainder of the paper is organized as follows. Section 2 introduces basic notions related to general binary relations and information systems. Section 3 introduces the concepts of entropy, joint entropy, conditional entropy, and mutual information for general binary relations and examines important properties. Section 4 discusses the potential applications for uncertainty measures of general binary relations. Section 5 concludes the study.

2. Basic notions

In this section, we review some basic notions related to general binary relations, information systems, and decision tables. For more detailed information, see the related content in Refs. [23,40].

Definition 2.1. Let $U = \{x_1, x_2, \cdots, x_n\}$ be a universe of discourse and R be a binary relation on U. For any $x \in U$, if $(x, x) \in R$, then R is referred to as reflexive. For any $x, y \in U$, if $(x, y) \in R$ implies $(y, x) \in R$, then R is referred to as symmetrical. For $x, y, z \in U$, if $(x, y) \in R$ and $(y, z) \in R$ implies $(x, z) \in R$, then R is referred to as transitive.

If $(x, y) \in R$ for any $x, y \in U$, then R is referred to as a universal relation. If $(x, x) \in R$ for any $x \in U$ and $(x, y) \notin R$ for any two $x, y \in U (x \neq y)$, then R is referred to as an identity relation. If R is reflexive, symmetrical, and transitive, then R is referred to as an equivalence relation.

Definition 2.2. Let R be a binary relation on U. For any $x \in U$, let

$$R(x) = \{y \in U | (x, y) \in R\}.$$

$R(x)$ is then called the successor neighborhood of x. Obviously, R: $U \to P(U)$ is a neighborhood operator, where $P(U)$ denotes the power set of U.

A binary relation R and its neighborhood operator R: $U \to P(U)$ are uniquely determined by each other. If R and S are two binary relations on U, then the following properties hold:

(1) $R \subseteq S \Leftrightarrow \forall x \in U, R(x) \subseteq S(x)$;
(2) $(R \cap S)(x) = R(x) \cap S(x)$;
(3) $(R \cup S)(x) = R(x) \cup S(x)$.

Let U be a nonempty set of samples $\{x_1, x_2, \cdots, x_n\}$ and A be a set of condition attributes for describing these samples in U, then the pair (U, A) is then called an information system. Given $B \subseteq A$ and R_B as a general binary relation induced by B, in the following discussion, we always assume that $R_B = \bigcap_{a \in B} R_a$.

Definition 2.3. Let (U, A) be an information system and $B \subseteq A$. Then, B is called a reduct of A if and only if (1) $R_B = R_A$ and (2) $R_B \neq R_{B-\{a\}}$ for any $a \in B$.

The first condition indicates that a reduct has the same discernibility as the whole attribute set. The second condition shows that no redundant attribute is present in the reduct.

Definition 2.4. The triple (U, A, D) is called a decision table, where A is a set of condition attributes that induce general binary relations and D is a decision attribute that induces an equivalence relation R_D on U.

Definition 2.5. Given a decision table (U, A, D), if $R_A \subseteq R_D$, then we say (U, A, D) is consistent. Let $B \subseteq A$, then B is called a relative reduct of A if and only if (1) $R_B \subseteq R_D$ and (2) $R_{B-\{a\}} \not\subset R_D$ for any $a \in B$.

From this definition, a relative reduct of A is a minimal subset of attributes that retain the same consistent classification as the entire set of attributes.

3. Information measures of general binary relations

In this section, we first define an equivalence relation based on the successor neighborhoods of samples and use the relation to divide the sample space into different parts. We then present the concepts of entropy and its variants for binary relations and discuss some basic properties.

Given a universe of discourse U, assume that R is a binary relation on U, R: $U \to P(U)$ is a neighborhood operator and $X \subseteq U$. Let $R(X) = \{R(x) | x \in X\}$. We then call $R(X)$ the neighborhood image of X.

The concept of classical entropy is of course based on probability distribution. To maintain the idea of classical entropy, we use an equivalence relation to group samples into different equivalence classes and employ them to define the entropy for general binary relations. We start by introducing the concept of equivalence relation as follows.

Definition 3.1. Given a universe U and a general binary relation R on U. Let

$$ER = \{(x_i, x_j) \in U \times U | R(x_i) = R(x_j)\}.$$

Obviously, ER is reflexive, symmetrical, and transitive. Hence, ER is an equivalence relation on U. The partition induced by ER is denoted as $U/ER = \{X_1, X_2, \cdots, X_r\}$, where $X_k = \{x_j \in U | R(x_i) = R(x_j)\}$, $x_i \in X_k$, $k = 1, 2, \cdots, r$.

The equivalence relation ER divides the universe U into mutually disjoint subsets $\{X_1, X_2, \cdots, X_r\}$. Here, we derive the following property about general binary relations

Proposition 3.1. Let R be a general binary relation on U, ER be the equivalence relation induced by R, and $U/ER = \{X_1, X_2, \cdots, X_r\}$. For any $x_i \in X_k (1 \le k \le r)$, we have (1) $R(X_k) = R(x_i)$ and (2) $X_k \subseteq R(x_i)$ if R is reflexive.

Proof. Straightforward.

Example 1. Let $U = \{x_1, x_2, x_3, x_4, x_5\}$ and R be one general binary relation on U, where