Information Sciences 212 (2012) 33–43

Contents lists available at SciVerse ScienceDirect

Information Sciences

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

Information-theoretic measures associated with rough set approximations

Ping Zhu [a,b,*], Qiaoyan Wen [b]

[a] School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China
[b] State Key Laboratory of Networking and Switching, Beijing University of Posts and Telecommunications, Beijing 100876, China

ARTICLE INFO

Article history:
Received 1 February 2011
Received in revised form 9 May 2012
Accepted 16 May 2012
Available online 24 May 2012

Keywords:
Rough set
Entropy
Co-entropy
Uncertainty
Granularity

ABSTRACT

In rough set theory, some information-theoretic measures of uncertainty and granularity have been proposed. A common feature of these measures is that they are only dependent on the partitions and the cardinality of a universe, which means that they are independent of the lower and upper approximations of rough sets. This seems somewhat unreasonable since the basic idea of rough set theory is to describe incomplete or inexact concepts by the lower and upper approximations. In light of this, we develop a new pair of information-theoretic entropy and co-entropy functions associated to partitions and approximations in this paper. Such functions are used to measure the uncertainty and granularity of an approximation space. After introducing the novel notions of entropy and co-entropy, we then examine their properties. In particular, we disclose the relationship of co-entropies between different universes. The theoretical development is accompanied by illustrative numerical examples.

1. Introduction

To handle incomplete or inexact knowledge in some information systems, Pawlak developed rough set theory in the early 1980s [21,22]. Since then, we have witnessed a systematic, world-wide growth of interest in rough set theory and its applications in many fields, such as granular computing, data mining, decision analysis, pattern recognition, and approximate reasoning [17,24,25,49,54,55].

The starting point of rough set theory in [21,22] is that the elements of a universe having the same description are indiscernible with respect to the available information. The indiscernibility is characterized by an equivalence relation in the way that two elements are equivalent if and only if they are indiscernible from each other. As is well known, any equivalence relation defined on a universe U determines a partition of U into a collection of equivalence classes (i.e., blocks): each class contains all and only the elements that are mutually equivalent among them. Any partition π of U represents a piece of knowledge about the elements of U which forms a classification, and so the equivalence class induced by π is interpreted as a granule of knowledge contained in (or supported by) π.

According to Pawlak's terminology expressed in [23], any subset X of a universe U is called a concept in U. If the concept X is a union of equivalence classes from π, then X is precise in π, otherwise X is vague. The basic idea of rough set theory consists in replacing incomplete or inexact concepts with a pair of precise concepts—its lower and upper approximations. To measure numerically these approximations, Pawlak [22] introduced two quantitative measures: accuracy and roughness. The accuracy of a concept is defined as the ratio of the cardinalities of the lower and upper approximations of the concept, while the roughness of the concept is calculated by subtracting the accuracy of the concept from 1. It should be stressed that this is an approach in which both the universe and the related partition are fixed once for all.

* Corresponding author at: School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China.
E-mail addresses: pzhubupt@gmail.com (P. Zhu), wqy@bupt.edu.cn (Q. Wen).

0020-0255/$ - see front matter
http://dx.doi.org/10.1016/j.ins.2012.05.014

One of the problems investigated by the rough set community is as follows: What happens when something changes? Two possible scenarios would be involved in this respect: The first one, which has been widely treated in the literature (see [2,44,50] and the bibliographies therein), is that the universe is fixed, but the involved partition changes according to some variations of the information available on the system (for instance, increasing of attributes or their reduction in information systems based on the same universe). The second one is that the universe changes, which necessarily leads to different partitions. Note that although there are some works dealing with objects entering/exiting a system and thus altering the universe (see, for example, [4,32,39]), the study from the perspective of uncertainty measures is initiated in the present paper.

For the first scenario, recall that there is a standard partial order "$\preceq$" defined on the collection $\Pi(U)$ of all partitions of a fixed universe U: For any $\pi,\sigma \in \Pi(U)$, $\pi \preceq \sigma$ if and only if for any $C \in \pi$, there exists $D \in \sigma$ such that $C \subseteq D$. With respect to this partial order, $\Pi(U)$ has a natural structure of complete lattice. The Pawlak roughness measure of the subsets of a universe is monotonically increasing, but it is not strictly monotonic. This means that there may exist two strictly ordered partitions $\pi \prec \sigma$ (i.e., $\pi \preceq \sigma$, but $\pi \neq \sigma$) and a subset A of U such that A has the same roughness measure with respect to π and σ. In order to avoid this unpleasant drawback, a large number of measures of partitions have been introduced [1,3,7,13,16,20,28,31,38,41–43,46,51]. Note also that the partial order relation plays an important role in the investigation of the monotonic behavior of various partition measures. As a result, the product of the Pawlak roughness measure with a partition measure turns out to be a strictly monotonically increasing measure of the subsets of U (see [50] and the bibliographies therein). Among the partition measures, there are several information-theoretic measures of uncertainty and granularity for rough sets [1,3,13,15,16,19,20,31,38,42], which are based upon the notion of entropy introduced by Shannon [33]. For more details, we refer the reader to the excellent survey papers [2,44].

It is worth noting that the information-theoretic measures mentioned above are only dependent on the sizes of equivalence classes (essentially, the underlying partitions) and the cardinality of a universe, independent of the lower and upper approximation operators. For example, in [7,20,38,43] the information entropy $H(\pi)$ of the partition $\pi = \{U_1, U_2, \ldots, U_k\}$ is defined as

$$H(\pi) = -\sum_{i=1}^{k} \frac{n_i}{n} \log \frac{n_i}{n},$$

where n_i is the cardinality of U_i and $n = \sum_{i=1}^{k} n_i$. It has been used as an average measure of knowledge granules in the approximation space $\langle U,\pi \rangle$. This seems somewhat unreasonable since the basic idea of rough set theory aims at describing incomplete or inexact concepts by the lower and upper approximations. In other words, the result of this description should rely on both the partition and the approximations. In light of this, we should pay more attention to the lower and upper approximation operators when considering various knowledge granulations (also, information granulations and granulation measures) in the framework of rough sets.

The previous observation motivates us to propose another information-theoretic entropy function to measure the uncertainty associated to a partition and approximation operators in the paper. More concretely, given a universe U with n elements and a partition π of U, we take count of the subsets of U described by every pair of lower and upper approximations. Assume that r_i, $1 \leqslant i \leqslant m$, is the number of subsets described by a rough set approximation (A_i, A_i') and every subset of U appears with the same probability. Then it follows that the rough set approximation (A_i, A_i') appears with the accumulative probability $r_i/2^n$ since the amount of all subsets of U is precisely 2^n. In this way, we obtain a probability distribution

$$P(\pi) = \left(\frac{r_1}{2^n}, \frac{r_2}{2^n}, \ldots, \frac{r_m}{2^n}\right).$$

It gives rise to an information entropy, say $\mathcal{H}(\pi)$, according to Shannon's information theory [33]. On the other hand, using the Hartley measure [10] we can also get a co-entropy $\mathcal{G}(\pi)$ by the probability distribution. It turns out that $\mathcal{H}(\pi) + \mathcal{G}(\pi) = n$.

After exploring some properties of the entropy and co-entropy, we discuss the relationships of co-entropies between different universes. In this case, we have to do with necessarily different partitions. Roughly speaking, the co-entropy monotonically increases when the partition becomes coarser. For example, the co-entropy of $\{\{1, 2\}, \{3, 4\}\}$ is greater than that of $\{\{1\}, \{2\}\}$. Formally, for any two approximation spaces $\langle U, \pi \rangle$ and $\langle V, \sigma \rangle$, we show that the co-entropy of π is not greater than that of σ if there is a monomorphism f from $\langle U, \pi \rangle$ to $\langle V, \sigma \rangle$. Here, by monomorphism we mean that f is an injective mapping from U to V and for any $C \in \pi$, there exists $D \in \sigma$ that includes the image of C under f. Furthermore, we clarify when the co-entropy is strictly monotonic as well.

The remainder of the paper is structured as follows. In Section 2, we briefly review some basics of Pawlak's rough set theory and two information-theoretic measures of uncertainty and granularity for rough sets in the literature. Section 3 is devoted to our novel notions of entropy and co-entropy and their properties. We address the relationship of co-entropies between different universes in Section 4 and conclude the paper in Section 5 with a brief discussion on the future research.