

Information granularity of soft sets

Yaya Liu^{1, a}, Keyun Qin^{2, b,*}

¹School of Mathematics, Southwest Jiaotong University, Chengdu 610031, Sichuan, PR China

²School of Mathematics, Southwest Jiaotong University, Chengdu 610031, Sichuan, PR China

^aemail: yayaliu@my.swjtu.edu.cn, ^bemail: qinkeyun@home.swjtu.edu.cn

Keywords: Soft set, Granular computing, Information granularity, Information entropy

Abstract. Soft set theory, proposed by Molodtsov, has been regarded as an effective mathematical tool for dealing with uncertainties. This paper is devoted to the study of uncertainty measures of soft sets in the framework of granular computing. The similarity relation induced by a soft set is proposed and the related fuzzy granular structure is constructed. The notions of information granularity and information entropy of a soft set are introduced. Some formulas are presented to calculate them.

Introduction

In 1999, Molodtsov [11] introduced the concept of soft sets, which can be seen as a new mathematical tool for dealing with uncertainties. In practice, a complicated object is usually described from different aspects. Soft set theory is closely related to fuzzy set theory and rough set theory [1,11]. Maji et al. [21] initiated the study on hybrid structures involving soft sets and fuzzy sets. Afterwards, many researchers have worked on this concept. Various kinds of extended fuzzy soft sets such as generalized fuzzy soft sets [27], intuitionistic fuzzy soft sets [22,25], and vague soft sets [4] were presented.

The study of uncertainty measures is an important topic for the theories to deal with uncertainty. Majumdar and Samanta [26,28] initiated the study of uncertainty measures of soft sets and fuzzy soft sets. Jiang [9] proposed some distance measures between intuitionistic fuzzy soft sets, and constructed some entropies on intuitionistic fuzzy soft sets and interval-valued fuzzy soft sets. Wang and Qu [3] introduced axiomatic definitions of entropy, similarity measure and distance measure for vague soft sets, and some formulas have also been put forward to calculate them. Based on fuzzy implication operators and fuzzy equivalences, Qin et al. [23] proposed a new category of inclusion measures and similarity measures for fuzzy soft sets. Liu et al. [20] proposed some entropies for fuzzy soft sets. We notice that the entropies presented in [3,9,20] are based on the entropies for fuzzy sets. If the fuzzy soft sets (resp. intuitionistic fuzzy soft sets, vague soft sets) degenerate to soft sets, the entropies will be zero (take minimum value). Thus the uncertainty degree of soft set cannot be described.

Granular computing (GrC) plays a fundamental role in fuzzy information processing. Information granularity and information entropy are two main approaches to measuring the uncertainty of a granular structure. The granular structures induced by equivalence relation, tolerance relation and fuzzy relation have been extensively studied in the framework of GrC [5 – 7,13 – 19,24,30]. This paper initiated a theoretical study on the information granularity of soft sets. The notions of information granularity and information entropy of a soft set are introduced to measure its uncertainty. Some formulas are proposed to calculate the information granularity and information entropy.

Overview of fuzzy sets and soft sets

The theory of fuzzy sets initiated by Zadeh [2] provides a framework for representing and processing vague concepts by allowing partial memberships.

Let U be a nonempty set, called universe. A fuzzy set μ on U is defined by a membership function

$\mu:U \rightarrow [0,1]$. For $x \in U$, the membership value $\mu(x)$ essentially specifies the degree to which x belongs to the fuzzy set μ . There are many different definitions for fuzzy set operations. With the min-max system proposed by Zadeh [2], fuzzy set intersection, union and complement are defined componentwise as follows:

$$(\mu \cap \nu)(x) = \mu(x) \wedge \nu(x),$$

$$(\mu \cup \nu)(x) = \mu(x) \vee \nu(x),$$

$$\mu^c(x) = 1 - \mu(x),$$

where μ, ν are fuzzy sets on U and $x \in U$. μ is called a fuzzy subset of ν if $\mu(x) \leq \nu(x)$ for any $x \in U$.

Let U be the universe set and E the set of all possible parameters under consideration with respect to U . Usually, parameters are attributes, characteristics, or properties of objects in U . (U, E) will be called a soft space. Molodtsov defined the notion of a soft set in the following way:

Definition 1 [11] A pair (F, A) is called a soft set over U , where $A \subseteq E$ and F is a mapping given by

$$F: A \rightarrow P(U)$$

(1)

In other words, a soft set over U is a parameterized family of subsets of U . A is called the parameter set of the soft set (F, A) . For $e \in A$, $F(e)$ may be considered as the set of e -approximate elements of (F, A) .

Definition 2 [29] (1) A soft set (F, A) over U is said to be a null soft set and denoted by \emptyset_A if $F(e) = \emptyset$ for each $e \in A$;

(2) A soft set (F, A) over U is said to be an absolute soft set and denoted by U_A if $F(e) = U$ for each $e \in A$.

Fuzzy information granularity

By using a fuzzy binary relation, objects are granulated into a set of fuzzy information granules [5,7], and a fuzzy granular structure is formed. Formally, let R be a fuzzy similarity relation on the universe $U = \{x_1, x_2, \dots, x_n\}$. R induces a fuzzy granular structure:

$$G_r(R) = (N_R(x_1), N_R(x_2), \dots, N_R(x_n)) \quad (2)$$

where $N_R(x_i)$ is the fuzzy similarity class of x_i and defined by $N_R(x_i)(x_j) = R(x_i, x_j)$ for every $x_i, x_j \in U$. We note that $N_R(x_i)$ is a fuzzy set on U . In what follows, we denote by $K(U)$ the set of all fuzzy granular structures on U .

To date, several forms of information granularity have been proposed according to various views and targets [12-14,17,24,30]. Among these studies, the partial order relation between granular structures plays a key role in characterizing the monotonicity of information granularity. Specifically, three kinds of partial order relations were mainly considered.

Definition 9 [24] Let $P, Q \in K(U)$ where $P = (S_p(x_1), S_p(x_2), \dots, S_p(x_n))$, $Q = (S_q(x_1), S_q(x_2), \dots, S_q(x_n))$.

(1) The binary relation \preceq_1 on $K(U)$ is defined as: $P \preceq_1 Q$ if $S_p(x_i) \subseteq Q_p(x_i)$ for each $x_i \in U$.

(2) The binary relation \preceq_2 on $K(U)$ is defined as: $P \preceq_2 Q$ if $|S_p(x_i)| \leq |Q_p(x_i)|$ for each $x_i \in U$.

(3) The binary relation \preceq on $K(U)$ is defined as: $P \preceq Q$ if there exists a bijection $\lambda: U \rightarrow U$ such that $|S_p(x_i)| \leq |S_q(\lambda(x_i))|$ for every $i \leq n$.

Where $|S_p(x_i)| = \sum_{j=1}^n S_p(x_i)(x_j)$ is the cardinality of $S_p(x_i)$. Clearly, the partial order relation \preceq_1 and \preceq_2 are special cases of \preceq . That is to say, if $P \preceq_1 Q$ (or $P \preceq_2 Q$), then $P \preceq Q$. P and Q are called equivalent and denoted by $P \approx Q$ if there exists a bijection $\lambda: U \rightarrow U$ such that $|S_p(x_i)| \leq |S_q(\lambda(x_i))|$ for every $i \leq n$. By $P \prec Q$ we mean $P \preceq Q$ and $\neg(P \approx Q)$.

Definition 10 [24] Let $K(U)$ be the collection of all fuzzy binary granular structures on U . If, for every $P \in K(U)$, there is a real number $G(P)$ with the following properties:

- (1) $G(P) \geq 0$ (nonnegative);
- (2) For every $P, Q \in K(U)$ if $P \approx Q$, then $G(P) = G(Q)$ (invariability);
- (3) For every $P, Q \in K(U)$, if $P \prec Q$, then $G(P) < G(Q)$ (monotonicity);

Then G is called a fuzzy information granularity.

Fuzzy information granularity has been extensively studied. Some formulas have been put forward from different point of views to calculate the information granularity [5,12,24].

The information granularity of a soft sets

Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set over U . From classification point of view, we think that every object can be classified by every parameter. Namely, $x \in F(e)$ implies that x is classified to $F(e)$ and $x \notin F(e)$ implies that x is classified to $U - F(e)$. In general, each classification approach of objects is based on a specific similarity relation on the universe. In what follows, we propose a similarity relation on the universe with respect to a soft set.

Based on rough set theory, a soft set $\mathfrak{S} \in (F, A)$ can be thought as an information system, and hence induces an indiscernibility relation R on the universe. Formally, for every $x, y \in U$, $(x, y) \in R$ if and only if $F(e)(x) = F(e)(y)$ for every $e \in A$, where $F(e)(x) = 1$ if $x \in F(e)$ and $F(e)(x) = 0$ otherwise. That is to say, $(x, y) \in R$ if and only if x and y are indiscernible with respect to every parameter $e \in A$ (i.e., $x \in F(e)$ and $y \in F(e)$, or $x \notin F(e)$ and $y \notin F(e)$).

In practice, however, it seems that considering the proportion of parameters that x and y are indiscernible with respect to them may lead to a better utilization of soft set approach. Thus we introduce the similarity degree $R_{\mathfrak{S}}(x, y)$ of objects x and y with respect to \mathfrak{S} as follows:

$$R_{\mathfrak{S}}(x, y) = \frac{1}{|A|} \cdot |\{e \in A; (\{x, y\} \subseteq F(e)) \vee (\{x, y\} \cap F(e) = \emptyset)\}| \quad (3)$$

Clearly, $R_{\mathfrak{S}} \in F(U \times U)$ is a fuzzy relation on the universe U . The following theorem is trivial.

Theorem 11 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set over U . Then

- (1) $R_{\mathfrak{S}}$ is reflexive, i.e., $R_{\mathfrak{S}}(x, x) = 1$ for all $x \in U$.
- (2) $R_{\mathfrak{S}}$ is symmetric, i.e., $R_{\mathfrak{S}}(x, y) = R_{\mathfrak{S}}(y, x)$ for all $x, y \in U$.
- (3) $R_{\mathfrak{S}}(x, y) = 1$ if and only if $(x, y) \in R$.
- (4) $R_{\mathfrak{S}}(x, y) = 0$ implies $(x, y) \notin R$.

Let $U = \{x_1, x_2, \dots, x_n\}$. The fuzzy granular structure $K(\mathfrak{S})$ associated with \mathfrak{S} is the vector:

$$K(\mathfrak{S}) = (N_{\mathfrak{S}}(x_1), N_{\mathfrak{S}}(x_2), \dots, N_{\mathfrak{S}}(x_n)) \quad (4)$$

where $N_{\mathfrak{S}}(x_i) \in F(U)$ ($1 \leq i \leq n$) is the fuzzy similarity class of x_i and is defined by:

$$N_{\mathfrak{S}}(x_i)(x_j) = R_{\mathfrak{S}}(x_i, x_j) \quad (5)$$

for each $x_j \in U$.

Based on [5,24], we propose the information granularity of a soft set. Let $G_1: S(U, E) \rightarrow [0, 1]$ be defined as: for each $\mathfrak{S} = (F, A) \in S(U, E)$,

$$G_1(\mathfrak{S}) = \frac{1}{n} \sum_{i=1}^n \frac{|N_{\mathfrak{S}}(x_i)|}{n} \quad (6)$$

where $|N_{\mathfrak{S}}(x_i)| = \sum_{j=1}^n N_{\mathfrak{S}}(x_i)(x_j)$ is the cardinality of $N_{\mathfrak{S}}(x_i)$.

Let $S_A(U, E) = \{(F, A); F : A \rightarrow P(U)\}$ be the set of all soft sets over U with parameter set A .

Theorem 12 G_1 is an information granularity in the sense that:

- (1) For every $\mathfrak{S} \in S_A(U, E)$, $G_1(\mathfrak{S}) \geq 0$.
- (2) For every $\mathfrak{S}, \delta \in S_A(U, E)$, if $K(\mathfrak{S}) \approx K(\delta)$, then $G_1(\mathfrak{S}) = G_1(\delta)$.
- (3) For every $\mathfrak{S}, \delta \in S_A(U, E)$, if $K(\mathfrak{S}) \prec K(\delta)$, then $G_1(\mathfrak{S}) < G_1(\delta)$.

Theorem 13 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set on U . Then

$$G_1(\mathfrak{S}) = 1 - \frac{2}{n^2 |A|} \sum_{e \in A} |F(e) \cap (U - F(e))| \quad (7)$$

Corollary 14 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set on U . Then

- (1) $G_1(\mathfrak{S}) = 1$ if and only if $F(e) = \emptyset$ or $F(e) = U$ for every $e \in A$.
- (2) $\frac{1}{2} \leq G_1(\mathfrak{S}) \leq 1$.

Remark 15 (1) If a soft set $\mathfrak{S} = (F, A)$ satisfies $F(e) = \emptyset$ or $F(e) = U$ for every $e \in A$, then any two objects are indiscernible with respect to every parameter. In this case, \mathfrak{S} does not contain any information for classification. By this corollary, the information granularity $G_1(\mathfrak{S})$ is maximum. Consequently, $G_1(\emptyset_A)$ and $G_1(U_A) (= 1)$ take maximum information granularity.

(2) If we hope to get a normal information granularity, we can define $G^*(\mathfrak{S}) = 2G_1(\mathfrak{S}) - 1$.

Clearly, we have $0 \leq G^*(\mathfrak{S}) \leq 1$

Qian et al. [24] developed an axiomatic constraint with a partial order relation which is defined in terms of the size of each fuzzy information granule. Based on [24], we propose the following definition.

Definition 16 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set on U . The information granularity $G_2(\mathfrak{S})$ of \mathfrak{S} is defined as

$$G_2(\mathfrak{S}) = - \sum_{i=1}^n \frac{1}{n} \log_2 \frac{1}{|N_{\mathfrak{S}}(x_i)|} \quad (8)$$

Theorem 17 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a deterministic soft set on U . Then

$$G_2(\mathfrak{S}) = \frac{1}{n} \sum_{e \in A} |F(e)| \log_2 \left(n - \frac{2}{|A|} |U - F(e)| \right) \quad (9)$$

The information entropy of a soft sets

Information granularity and information entropy are two main approaches to measuring the uncertainty of a granular structure. Shannon's entropy of a equivalence granular structure is a measure of information uncertainty. It is defined by the probability of each equivalence information granule. It has been extended to dealing with tolerance granular structure and fuzzy granular structure [5,10,15,24]. Based on [24], we propose the information entropy of a soft set.

Definition 18 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set on U . The information entropy $E_1(\mathfrak{S})$ of \mathfrak{S} is defined as

$$E_1(\mathfrak{S}) = \sum_{i=1}^n \frac{1}{n} \left(1 - \frac{|N_{\mathfrak{S}}(x_i)|}{n} \right) \quad (10)$$

Theorem 19 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set on U . Then

$$E_1(\mathfrak{S}) = \frac{2}{n^2 |A|} \sum_{e \in A} |F(e)| |U - F(e)| \quad (11)$$

Definition 20 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set on U . The information entropy

$E_2(\mathfrak{S})$ of \mathfrak{S} is defined as

$$E_2(\mathfrak{S}) = -\frac{1}{n} \sum_{i=1}^n \log_2 \frac{|N_{\mathfrak{S}}(x_i)|}{n} \quad (12)$$

Theorem 21 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a deterministic soft set on U . Then

$$E_2(\mathfrak{S}) = \log_2 n - \frac{1}{n} \sum_{e \in A} |F(e)| \log_2 \left(n - \frac{2}{|A|} |U - F(e)| \right) \quad (13)$$

Corollary 22 Let (U, E) be a soft space and $\mathfrak{S} = (F, A)$ a soft set on U . Then

- (1) $G_1(\mathfrak{S}) + E_1(\mathfrak{S}) = 1$.
- (2) $G_2(\mathfrak{S}) + E_2(\mathfrak{S}) = \log_2 n$.

Conclusion

This paper is devoted to the study of uncertainty measures of soft sets in the framework of granular computing. The similarity relation induced by a soft set is proposed and the related fuzzy granular structure is constructed. Based on the related works on fuzzy granular structures, the notions of information granularity and information entropy of a soft set are introduced and investigated.

Acknowledgement

This work has been supported by the National Natural Science Foundation of China (Grant No. 61473239, 61372187) and The Fundamental Research Funds for the Central Universities of China (Grant No.2682014ZT28).

References

- [1] Hacı Aktaş and Naim Çağman. Soft sets and soft groups[J]. Information Sciences. 2007:177:2726-2735.
- [2] Lotfi A. Zadeh. Fuzzy sets[J]. Information and Control. 1965:8:338-353.
- [3] Chang Wang and Anjing Qu. Entropy, similarity measure and distance measure of vague soft sets and their relations[J]. Information Sciences. 2013:244:92-106.
- [4] Wei Xu, et al. Vague soft sets and their properties[J]. Computers and Mathematics with Applications. 2010:59:787-794.
- [5] Qinghua Hu, et al. Fuzzy probabilistic approximation spaces and their information measures[J]. IEEE Transactions on Fuzzy Systems. 2006:14(2):253-264.
- [6] Qinghua Hu, et al. EROS: Ensemble rough subspaces[J]. Pattern Recognition. 2007(40)3728-3739.
- [7] Jensen Richard, and Qiang Shen. New approaches to fuzzy-rough feature selection[J]. IEEE Transactions on Fuzzy Systems. 2009:17(4):824-838.
- [8] Yuncheng Jiang, et al. Interval-valued intuitionistic fuzzy soft sets and their properties[J], Computers and Mathematics with Applications. 2010:60:906-918.
- [9] Yuncheng Jiang, et al. Entropy on intuitionistic fuzzy soft sets and on interval-valued fuzzy soft sets[J]. Information Sciences. 2013:240:95-114.
- [10] Claude Elwood Shannon. The mathematical theory of communication[J]. The Bell System

- Technical Journal. 1948:27(3-4):373-423, 623-656.
- [11] Dmitriy Molodtsov. Soft set theory-First results[J]. Computers and Mathematics with Applications. 1999:37:19-31.
- [12] Jiye Liang, et al. A new method for measuring uncertainty and fuzziness in rough set theory[J]. International Journal of General Systems. 2002:31(4):331-342.
- [13] Jiye Liang and Zhongzhi Shi. The information entropy, rough entropy and knowledge granulation in rough set theory[J]. International Journal of Uncertainty, Fuzziness Knowledge-Based Systems. 2004:12(1):37-46.
- [14] Jiye Liang, et al. The information entropy, rough entropy and knowledge granulation in incomplete information system[J]. International Journal of General Systems. 2006:35(6):641-654.
- [15] Tsau Young Lin. Neighborhood systems and approximation in database and knowledge base systems[C]. in: Proc. 4th Int. Symp. Methodol. Intell. Syst., Charlotte, NC. 1989, 75-86.
- [16] Tsau Young Lin. Granular computing on binary relations I: Data mining and neighborhood systems[C]. in: Rough Sets In Knowledge Discovery, A.Skowron and L.Polkowski, Eds. New York: Physica-Verlag. 1998, 107-120.
- [17] Tsau Young Lin. Granular computing on binary relations II: Rough set representations and belief functions[C]. in: Rough Sets in Knowledge Discovery, A.Skowron and L.Polkowski, Eds. New York: Physica-Verlag. 1998, 121-140.
- [18] Tsau Young Lin. Granular computing I: The concept of granulation and its formal model[J]. International Journal of Granular Computing, Rough Sets and Intelligent Systems. 2009:1(1):21-42.
- [19] Tsau Young Lin. Granular computing: practices, theories and future directions[C], in: Encyclopedia of Complexity and Systems Science, R.A.Meyers Ed., Berlin, Heidelberg: Springer. 2009, 4339-4355.
- [20] Zhicai Liu, et al. Similarity measure and entropy of fuzzy soft sets[J], The Scientific World Journal. 2014(4-5):161607-161607.
- [21] Pabitra Kumar Maji, et al. Fuzzy soft sets[J], The Journal of Fuzzy Mathematics. 2001:9(3):589-602.
- [22] Pabitra Kumar Maji, et al. Intuitionistic fuzzy soft sets[J], The Journal of Fuzzy Mathematics. 2001:9(3):677-692.
- [23] Keyun Qin, et al. Some uncertainty measures of fuzzy soft sets[C], in: Proceedings of 2013 Asia-Pacific Computational Intelligence and Information Technology, D.Zhong and A.Hu Eds. Shanghai: DEStech Publications. 2013, 245-251.
- [24] Yuhua Qian, et al. Information granularity in fuzzy binary GrC model[J], IEEE Transactions on Fuzzy Systems. 2011:19(2):253-264.
- [25] Pabitra Kumar Maji, et al. On intuitionistic fuzzy soft sets[J], The Journal of Fuzzy Mathematics. 2004:12(3):669-683.
- [26] Pinaki Majumdar and Syamal Kumar Samanta. Similarity measure of soft sets[J], New Mathematics and Natural Computation. 2008:4(1):1-12.
- [27] Pinaki Majumdar and Syamal Kumar Samanta. Generalised fuzzy soft sets[J], Computers and Mathematics with Applications. 2010:59:1425-1432.
- [28] Pinaki Majumdar and Syamal Kumar Samanta. On similarity measures of fuzzy soft sets[J], International Journal of Advance in Soft Computing and its Application. 2011:3(2):1-8.
- [29] Pabitra Kumar Maji, et al. Soft set theory[J], Computers and Mathematics with Applications. 2003:45:555-562.
- [30] Yuhua Qian, et al. Incomplete multigranulation rough set[J], IEEE Transactions on Systems, Man, and Cybernetics, Part A, Systems and humans. 2010:40(2):420-431.