



Two-way Concept-Cognitive Learning with Multi-source Fuzzy Context

Xiaoyan Zhang¹ · Doudou Guo¹ · Weihua Xu¹

Received: 23 November 2021 / Accepted: 2 January 2023 / Published online: 22 February 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Concepts learning is the most fundamental unit in the process of human cognition in philosophy. Granularity is one of the fundamental concepts of human cognition. The combination of granular computing and concept learning is critical in the cognitive process. Meanwhile, efficiently and accurately using the information collected from different sources is the focus of data mining in the contemporary. Hence, how to sufficiently learn concepts under a multi-sources context is an essential concern in the field of cognition. This paper offers a new thought for two-way concept-cognitive learning based on granular computing in multi-source fuzzy decision tables. Firstly, based on the best possible guarantee of the classification ability, original information from different sources is fused by conditional entropy, which is the kind of multi-source fusion method (i.e., CE-fusion). Secondly, we learn concepts from a given object set, attribute set, or pair of object and attribute sets in the fused information table, and these three types of concept learning algorithms are designed. This analysis shows that two-way concept learning based on multi-source information fusion is a suitable method of multi-source concept learning. Some examples are valuable for applying these theories to deal with practical issues. Our work will provide a convenient novel tool for researching concept-cognitive learning methods with multi-source fuzzy context.

Keywords Concept-cognitive learning · Two-way learning · Granular computing · Multi-source fuzzy context · Formal concept analysis

Introduction

The basic concept and scientific methodology of cognitive science are built on modern scientific analysis and engineering experiments to study cognition and intelligence. Up to now, cognitive computing has been viewed as the development of computer systems modeled on the human brain. As is well known, concepts are the most fundamental units of human cognition, and concept learning is the most fundamental unit in the process of human cognition in philosophy [1, 2]. The critical mission of conceptual knowledge presenting and processing is to obtain exact and accurate concepts from various aspects. Concept-cognitive learning

(CCL), the idea of cognitive through concept formation and learning to reveal the systematic law of the human brain, is an effective cognitive mechanism [1, 3–5]. It should be noted that CCL will be an essential viewpoint for us to carry out cognitive science research in the current article.

As an emerging paradigm of intelligent computing methodologies, cognitive computing has the characteristic of integrating past experiences into itself. So it is a beneficial idea to learn concepts from the cognitive viewpoint. Note that concepts can be characterized by their extent and intent, which can be determined by each other [6–8]. The extent of the concept is the scope of application, which the object set can express that the concept denotes. The intent of concept is the unique attribute reflected by the concept, which the attribute can express a set that a concept connotes [8–10]. So concepts can be learned from two aspects of intent and extent. As an emerging paradigm of intelligent computing methodologies, cognitive computing has the characteristic of integrating past experiences into itself [3, 11, 12]. So it is a good idea to learn concepts from the cognitive viewpoint. Now, concept-cognitive learning, as the development of computer systems modeled on the human brain, is widely

✉ Doudou Guo
doudou876517690@126.com

Xiaoyan Zhang
zxy19790915@163.com

Weihua Xu
chxuwh@gmail.com

¹ College of Artificial Intelligence, Southwest University, Chongqing 400715, China

concerned [1, 13–15]. Note that authors [5, 16] firstly investigated learning concepts from the unknown through a pair of cognitive operators (i.e., extent-intent and intent-extent) to simulate human thought processes. Hence, two-way learning has become the primary method and theory basis for learning the concept from the multi-source information table in this paper.

Currently, big data has opened up a whole new era for cognitive science [17, 18]. As an essential basis to support cognitive science, data and knowledge fusion driven is related to the formulation of man-machine intelligence. One of the most urgent things is how to use data from different sources to discover effective knowledge. The deficiency of the single data can be made up via integrating data from different sources to achieve the mutual complement and confirmation of a variety of data sources. The means of multi-source fusion expand the application range of data and improve the accuracy of the analysis. Therefore, taking full advantage of multi-source information is crucial to learn concepts and knowledge. This paper emphasizes the concept-cognitive learning of multi-source fuzzy decision tables. It is also important to note that effective concepts may not be learned through existing concept learning methods in a single decision table. Thus, concept learning under multi-source information is worth paying attention to.

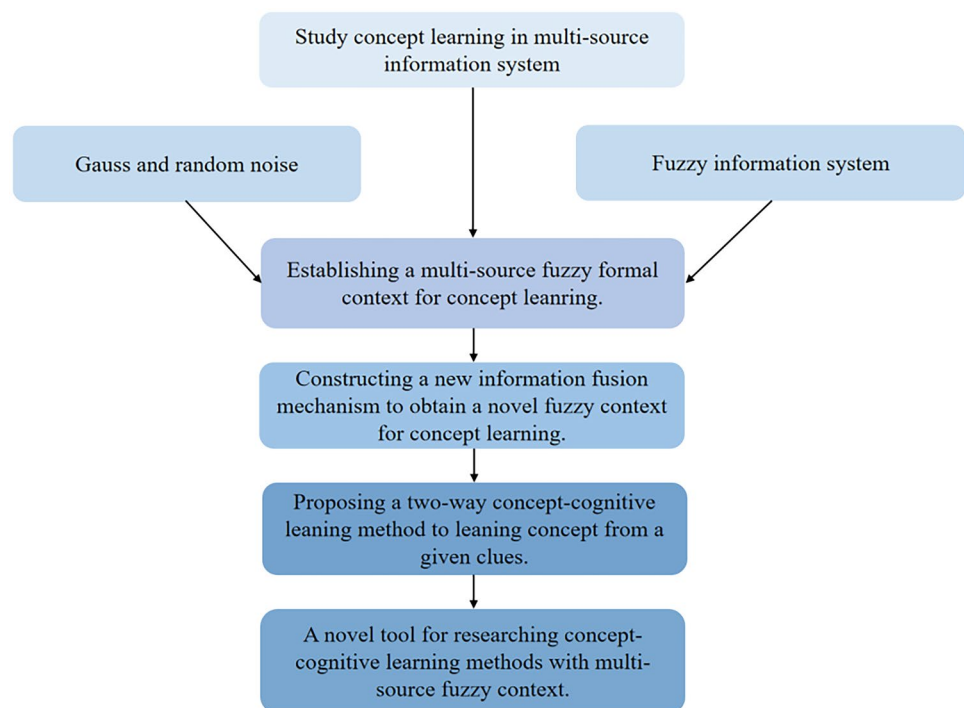
From a philosophical viewpoint, granular computing (GrC) is a structured way of thinking. From the application viewpoint, GrC is a general method for solving structured problems. From the calculation viewpoint, GrC is a typical method of information processing. At present,

there are many studies on granular computing in references [19–22]. Information granules are formalized in many different ways. They can be expressed in terms of the set, fuzzy set, rough set, formal concept, etc. Formal concept analysis (FCA) first proposed by Wille [8] is a popular human-centered tool for knowledge discovery, data mining, and bi-clustering and has widely been applied to lots of fields [14, 15, 23]. Note that these certain concept structures or lattices establish rigorous mathematical models and provide a formal semantics for data analysis in practice [24, 25]. In other words, meanings of real-world, concrete entities can be represented, and these certain concept structures can embody the semantics of abstract subjects. All in all, learning concepts (sometimes including their corresponding structure) have been investigated from various aspects. In addition, considering the computational complexity and the uncertainty of concept learning in multi-source decision tables, this paper focuses on two-way concept-cognitive learning with multi-source information fusion by granular computing.

We proposed a two-way concept-cognitive learning approach to learn fuzzy concept from multi-source fuzzy decision tables. The block diagram of steps of the proposed approach is shown in Fig. 1. The main contributions of this paper are as follows:

1. We discuss a new thought of concept learning based on the multi-source decision table by connecting two-way learning, granular computing to multi-source information fusion theory. Furthermore, it attempts to construct

Fig. 1 Block diagram of steps of the proposed approach



a new concept learning method from a multi-source information fusion.

- To take full advantage of multi-source information to learn concepts, we build two kinds of multi-source fuzzy decision fusion mechanisms, including conditional entropy fusion and mean-fusion. Moreover, some examples are used to verify the effectiveness of the two fusion mechanisms.
- Compared with other CCL models, the two-way concept-cognitive learning method emphasizes learning concepts from clues. Therefore, the mechanic two-way concept-cognitive learning for given specific tasks or cues is more advantageous than the current many CCL model.

The remainder of this paper is organized as follows. The “[Preliminaries](#)” section briefly reviews some basic notions about granule computing and formal concept analysis. The “[Two-way Concept-Cognitive Learning](#)” section discusses a cognitive concept mechanism of transforming any information granule into a sufficient and necessary information granule. The definition of the multi-source fuzzy decision table and multi-source fusion is constructed to learn the concept based on entropy theory in the “[Concept Learning in Multi-source Fuzzy Decision Tables](#)” section. Furthermore, the “[Conclusions](#)” section concluded with a discussion for further work.

Preliminaries

This section briefly reviews some basic notions related to (1) fuzzy set, (2) an uncertainly measure in the rough set, and (3) fuzzy formal context. More detailed descriptions can reference the literatures [6, 26–28].

Fuzzy Set

Fuzzy set first proposed by Zadeh [28] attaches great importance to the idea of partial membership, which departs from the dichotomy. Fuzzy set theory is a generalization of the classical set theory. Let U be a nonempty finite set, and a fuzzy set \tilde{X} of U can be expressed as:

$$\tilde{X} = \{ \langle x, \mu_{\tilde{X}}(x) \rangle \mid x \in U \},$$

where $\mu_{\tilde{X}}(x): U \rightarrow [0, 1]$, $\mu_{\tilde{X}}(x)$ is called the membership degree of the object $x \in U$ with respect to \tilde{X} . Let $\mathcal{F}(U)$ customarily denote all fuzzy sets in the universe U . Given two fuzzy sets $\tilde{X}_1, \tilde{X}_2 \in \mathcal{F}(U)$, for any $x \in U$, $\mu_{\tilde{X}_1}(x) \leq \mu_{\tilde{X}_2}(x)$ if and only if $\tilde{X}_1 \subseteq \tilde{X}_2$; $\mu_{\tilde{X}_1}(x) = \mu_{\tilde{X}_2}(x)$ if and only if $\tilde{X}_1 \subseteq \tilde{X}_2$ and $\tilde{X}_2 \subseteq \tilde{X}_1$. Let $\sim \tilde{X}$ represent the complement set of \tilde{X} , $\tilde{X}_1 \cap \tilde{X}_2$ represents the intersection of \tilde{X}_1 and \tilde{X}_2 , and $\tilde{X}_1 \cup \tilde{X}_2$ represents the union of \tilde{X}_1 and \tilde{X}_2 , and specific expressions are as follows:

$$\begin{aligned} \sim \tilde{X} &= \{ \langle x, 1 - \mu_{\tilde{X}}(x) \rangle \mid x \in U \}, \\ \tilde{X}_1 \cap \tilde{X}_2 &= \{ \langle x, \wedge \{ \mu_{\tilde{X}_1}(x), \mu_{\tilde{X}_2}(x) \} \rangle \mid x \in U \}, \\ \tilde{X}_1 \cup \tilde{X}_2 &= \{ \langle x, \vee \{ \mu_{\tilde{X}_1}(x), \mu_{\tilde{X}_2}(x) \} \rangle \mid x \in U \}. \end{aligned}$$

A decision system is a quadruple $I = (U, A, V, f)$, where U is a nonempty finite universe; $A = C \cup D$ is the union of condition attribute set C and decision attribute set D , and $C \cap D = \emptyset$; V is the union of attribute domains, i.e., $V = \cup_{a \in A} V_a$; $f: U \times A \rightarrow V$ is an information function, i.e., $\forall a \in A, x \in U$, that $f(x, a) \in V_a$, where $f(x, a)$ is the value of the object x under the attribute a . Generally, let $D = \{d\}$. Particularly, if the value range of V is from 0 to 1, then the decision system $I = (U, A, V, f)$ is called a fuzzy decision system which can be denoted by $\tilde{I} = (U, A, V, f)$.

An Uncertainty Measure in Rough Set

Rough set is one of effective mathematical tools for data analysis and knowledge discovery. Let U be a nonempty finite universe and R be an equivalence relation of $U \times U$. The equivalence relation R induces a partition of U , denoted by $U/R = \{[x]_R \mid x \in U\}$, where $[x]_R$ represents the equivalence class of x with regard to R . Then, (U, R) is called the Pawlak approximation space. For an arbitrary subset X of U , the lower and upper approximations of X are defined as follows:

$$\begin{aligned} \underline{R}(X) &= \{x \in U \mid [x]_R \cap X \neq \emptyset\} = \cup \{[x]_R \mid [x]_R \cap X \neq \emptyset\}, \\ \overline{R}(X) &= \{x \in U \mid [x]_R \subseteq X\} = \cup \{[x]_R \mid [x]_R \subseteq X\}. \end{aligned}$$

And $pos(X) = \underline{R}(X)$, $neg(X) = \sim \overline{R}(X)$, $bnd(X) = \overline{R}(X) - \underline{R}(X)$ are called the positive region, negative region, and boundary region of X , respectively. Objects definitely and not definitely contained in the set X form positive region $pos(X)$ and negative region $neg(X)$. Objects that may be contained in the set X constitute boundary region $bnd(X)$.

The uncertainty measure is an important direction in rough set theory. The approximation precision proposed by Pawlak raised the proportion of correct classification by an equivalence relation. Let $I = (U, A, V, f)$ be a decision system and $U/D = \{D_1, D_2, \dots, D_m\}$ be a classification of universe U . For an arbitrary attribute subset R of C , the R -lower and R -upper approximations of U/D are defined as

$$\underline{R}(U/D) = \underline{R}(D_1) \cup \underline{R}(D_2) \cup \dots \cup \underline{R}(D_m),$$

$$\overline{R}(U/D) = \overline{R}(D_1) \cup \overline{R}(D_2) \cup \dots \cup \overline{R}(D_m).$$

The approximation precision and the corresponding approximation roughness of U/D by R are defined as

$$\alpha_R(U/D) = \frac{\sum_{Y_i \in U/D} |R(Y_i)|}{\sum_{Y_i \in U/D} |\bar{R}(Y_i)|}$$

$$Roughness_R(U/D) = 1 - \alpha_R(U/D).$$

Fuzzy Formal Context

Formal concept analysis (FCA) proposed by Wille [8] provides comprehensive knowledge about a dataset called a formal context. The formal context is a triplet of an object set, an attribute set, and an information function about objects and attributes whose domain consists of 0 and 1. The formal context is fuzzy when the domain is the natural interval [0, 1]. Therefore, fuzzy information systems can be analyzed from the perspective of formal contexts. That is to say, fuzzy formal context is a triple (U, A, \tilde{I}) , where $U = \{x_1, x_2, \dots, x_n\}$ is an object set and $A = \{a_1, a_2, \dots, a_m\}$ is an attribute set and \tilde{I} is a fuzzy binary relation from U to A , $\tilde{I} = \{ \langle (x, a), \mu_{\tilde{I}}(x, a) \rangle \mid (x, a) \in U \times A \}$, $\mu_{\tilde{I}}(x, a) : U \times A \rightarrow [0, 1]$.

The cognitive mechanism of forming concepts can be described as follows. Let $X(X \subseteq U)$ be an object set and $B(B \subseteq A)$ be an attribute set. A pair of operators firstly be defined, namely

$$X^* = \tilde{A} = \{ \langle a, \mu_{\tilde{I}}(x) \rangle \mid a \in A \},$$

$$\tilde{B}^* = \{ x \in U \mid \tilde{I}(x, b) \geq \tilde{B}(b), b \in B \}.$$

where $\tilde{A}(a) = \bigwedge_{x \in X} \mu_{\tilde{I}}(x, a)$, $\tilde{I}(x, b) = \mu_{\tilde{I}}(x, b) \in \nu$ and $\nu = \{ \tilde{I}(x, a), x \in U, a \in A \}$. Particularly, we rule $\emptyset^* = \tilde{A} = \{ \langle a, 0 \rangle \mid a \in A \}$. From the perspective of philosophy, a concept consists of two parts: extent Y which is a set of objects and intent C which is a set of attributes. In general, the more objects a concept denotes, the less attributes it connotes, and vice versa. That is to say, $X_1 \subseteq X_2 \Rightarrow X_2^* \subseteq X_1^*$ and $\tilde{B}_1 \subseteq \tilde{B}_2 \Rightarrow \tilde{B}_2^* \subseteq \tilde{B}_1^*$, where X_1^* and X_2^* denote the corresponding intents of X_1 and X_2 and \tilde{B}_1^* and \tilde{B}_2^* denote the corresponding extents of \tilde{B}_1 and \tilde{B}_2 . From the perspective of cognitive psychology, the perception of the whole is more than the integration of perceptions of its parts. That is to say, $(X_1 \cup X_2)^* \supseteq X_1^* \cap X_2^*$ and $(\tilde{B}_1 \cup \tilde{B}_2)^* \supseteq \tilde{B}_1^* \cap \tilde{B}_2^*$. Moreover, the above two operators have the following properties:

1. $X_1 \subseteq X_2 \Rightarrow X_2^* \subseteq X_1^*, \tilde{B}_1 \subseteq \tilde{B}_2 \Rightarrow \tilde{B}_2^* \subseteq \tilde{B}_1^*$
2. $X \subseteq X^{**}, \tilde{B} \subseteq \tilde{B}^{**}$
3. $X^* = X^{***}, \tilde{B}^* = \tilde{B}^{***}$
4. $X \subseteq \tilde{B}^* \Leftrightarrow \tilde{B} \subseteq X^*$
5. $(X_1 \cup X_2)^* = X_1^* \cap X_2^*, (\tilde{B}_1 \cup \tilde{B}_2)^* = \tilde{B}_1^* \cap \tilde{B}_2^*$

$$6. (X_1 \cap X_2)^* \supseteq X_1^* \cap X_2^*, (\tilde{B}_1 \cap \tilde{B}_2)^* \supseteq \tilde{B}_1^* \cap \tilde{B}_2^*$$

A pair (X, \tilde{B}) is called a fuzzy concept, if $X^* = \tilde{B}$ and $X = \tilde{B}^*$, for $X \subseteq U, B \subseteq A$. X and \tilde{B} are called the extent and intent of (X, \tilde{B}) , respectively. It is clear that both (X^{**}, X^*) and $(\tilde{B}^*, \tilde{B}^{**})$ are fuzzy concepts.

In order to more clearly explain the concept of learning process, we introduce a fuzzy formal context and detailed information is shown in Table 1. The universe is $U = \{x_1, x_2, x_3, x_4\}$ and the attribute set is $A = \{a, b, c, d\}$.

According to intuitive perception and attention, we can obtain the result of the object-oriented and attribute-oriented operators, shown in Tables 2 and 3.

When $X^* = \tilde{B}$ and $X = \tilde{B}^*$, the pair (X, \tilde{B}) is called a fuzzy concept. According to the definition of the fuzzy concept and the result of two operators, concepts of the formal context introduced can be obtained as follows: $(\emptyset, \{ \langle a, 1 \rangle, \langle b, 1 \rangle, \langle c, 1 \rangle, \langle d, 1 \rangle \})$, $(\{x_1\}, \{ \langle a, 0.6 \rangle, \langle b, 0.2 \rangle, \langle c, 0.5 \rangle, \langle d, 0.3 \rangle \})$, $(\{x_2\}, \{ \langle a, 0.3 \rangle, \langle b, 0.3 \rangle, \langle c, 0.7 \rangle, \langle d, 0.5 \rangle \})$, $(\{x_3\}, \{ \langle a, 0.7 \rangle, \langle b, 0.6 \rangle, \langle c, 0.2 \rangle, \langle d, 0.9 \rangle \})$, $(\{x_4\}, \{ \langle a, 0.2 \rangle, \langle b, 0.4 \rangle, \langle c, 0.4 \rangle, \langle d, 0.7 \rangle \})$, $(\{x_1, x_2\}, \{ \langle a, 0.3 \rangle, \langle b, 0.2 \rangle, \langle c, 0.5 \rangle, \langle d, 0.3 \rangle \})$, $(\{x_1, x_3\}, \{ \langle a, 0.6 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \})$, $(\{x_2, x_3\}, \{ \langle a, 0.3 \rangle, \langle b, 0.3 \rangle, \langle c, 0.2 \rangle, \langle d, 0.5 \rangle \})$, $(\{x_2, x_4\}, \{ \langle a, 0.2 \rangle, \langle b, 0.3 \rangle, \langle c, 0.4 \rangle, \langle d, 0.5 \rangle \})$, $(\{x_3, x_4\}, \{ \langle a, 0.2 \rangle, \langle b, 0.4 \rangle, \langle c, 0.2 \rangle, \langle d, 0.7 \rangle \})$, $(\{x_1, x_2, x_3\}, \{ \langle a, 0.3 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \})$, $(\{x_1, x_2, x_4\}, \{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.4 \rangle, \langle d, 0.3 \rangle \})$, $(\{x_2, x_3, x_4\}, \{ \langle a, 0.2 \rangle, \langle b, 0.3 \rangle, \langle c, 0.2 \rangle, \langle d, 0.5 \rangle \})$, $(\{x_1, x_2, x_3, x_4\}, \{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \})$.

Two-way Concept-Cognitive Learning

For a fuzzy formal context (U, A, \tilde{I}) , 2^U and 2^A be the power sets of U and A , respectively. The $\tilde{\mathcal{F}} : 2^U \rightarrow 2^A$ and $\mathcal{P} : 2^A \rightarrow 2^U$ are considered as a pair of two-way learning operators, and they are abbreviated as $\tilde{\mathcal{F}}$ and \mathcal{P} , respectively. $L_1 = P(U)$ and $\tilde{L}_2 = P(A)$ be complete lattices and fuzzy complete lattices, respectively.

Let L_1 and \tilde{L}_2 be a pair of complete lattices, for any $X_1, X_2 \in L_1, \tilde{\mathcal{F}} : L_1 \rightarrow \tilde{L}_2$ is an extent-intent cognitive operator if $\tilde{\mathcal{F}}$ satisfies the following:

Table 1 A fuzzy formal context

U	a	b	c	d
x_1	0.6	0.2	0.5	0.3
x_2	0.3	0.3	0.7	0.5
x_3	0.7	0.6	0.2	0.9
x_4	0.2	0.4	0.4	0.7

Table 2 Object-oriented operator

$\emptyset^* = \{ \langle a, 1 \rangle, \langle b, 1 \rangle, \langle c, 1 \rangle, \langle d, 1 \rangle \}$	$\{x_2, x_3\}^* = \{ \langle a, 0.3 \rangle, \langle b, 0.3 \rangle, \langle c, 0.2 \rangle, \langle d, 0.5 \rangle \}$
$\{x_1\}^* = \{ \langle a, 0.6 \rangle, \langle b, 0.2 \rangle, \langle c, 0.5 \rangle, \langle d, 0.3 \rangle \}$	$\{x_2, x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.3 \rangle, \langle c, 0.4 \rangle, \langle d, 0.5 \rangle \}$
$\{x_2\}^* = \{ \langle a, 0.3 \rangle, \langle b, 0.3 \rangle, \langle c, 0.7 \rangle, \langle d, 0.5 \rangle \}$	$\{x_3, x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.4 \rangle, \langle c, 0.2 \rangle, \langle d, 0.7 \rangle \}$
$\{x_3\}^* = \{ \langle a, 0.7 \rangle, \langle b, 0.6 \rangle, \langle c, 0.2 \rangle, \langle d, 0.9 \rangle \}$	$\{x_1, x_2, x_3\}^* = \{ \langle a, 0.3 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}$
$\{x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.4 \rangle, \langle c, 0.4 \rangle, \langle d, 0.7 \rangle \}$	$\{x_1, x_2, x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.4 \rangle, \langle d, 0.3 \rangle \}$
$\{x_1, x_2\}^* = \{ \langle a, 0.3 \rangle, \langle b, 0.2 \rangle, \langle c, 0.5 \rangle, \langle d, 0.3 \rangle \}$	$\{x_1, x_3, x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}$
$\{x_1, x_3\}^* = \{ \langle a, 0.6 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}$	$\{x_2, x_3, x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.3 \rangle, \langle c, 0.2 \rangle, \langle d, 0.5 \rangle \}$
$\{x_1, x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.4 \rangle, \langle d, 0.3 \rangle \}$	$\{x_1, x_2, x_3, x_4\}^* = \{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}$

- $\tilde{\mathcal{F}}(0_{L_1}) = 1_{\tilde{L}_2}, \tilde{\mathcal{F}}(1_{L_1}) = 0_{\tilde{L}_2}$,
- $\tilde{\mathcal{F}}(X_1 \vee X_2) = \tilde{\mathcal{F}}(X_1) \wedge \tilde{\mathcal{F}}(X_2)$.

Similarly, for any $\tilde{B}_1, \tilde{B}_2 \in \tilde{L}_2, \mathcal{P} : L_1 \rightarrow \tilde{L}_2$ is an extent-intent cognitive operator if

- $\mathcal{P}(0_{L_1}) = 1_{\tilde{L}_2}, \mathcal{P}(1_{L_1}) = 0_{\tilde{L}_2}$,
- $\mathcal{P}(\tilde{B}_1 \vee \tilde{B}_2) = \mathcal{P}(\tilde{B}_1) \wedge \mathcal{P}(\tilde{B}_2)$.

where $0_{\tilde{L}}$ and $1_{\tilde{L}}$ are zero and the unit element, respectively.

Definition 1 Let (U, A, \tilde{I}) be a fuzzy formal context, $L_1 = P(U)$ and $\tilde{L}_2 = P(\tilde{A})$ be two complete lattices, L and \mathcal{P} be two cognitive operators (i.e., $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$ is a cognitive system). For any $X \in L_1, \tilde{B} \in \tilde{L}_2$, denote

$$\tilde{\mathcal{G}}_1 = \{ (X, \tilde{B} | \tilde{B} \leq \tilde{\mathcal{F}}(X), X \leq \mathcal{P}(\tilde{B}) \},$$

$$\tilde{\mathcal{G}}_2 = \{ (X, \tilde{B} | \tilde{\mathcal{F}}(X) \leq \tilde{B}, \mathcal{P}(\tilde{B}) \leq X \}.$$

- If $(X, \tilde{B}) \in \mathcal{G}_1$, then (X, \tilde{B}) is a necessary fuzzy information granule and \tilde{B} is necessary fuzzy attribute of X .
- If $(X, \tilde{B}) \in \mathcal{G}_2$, then (X, \tilde{B}) is a sufficient fuzzy information granule and \tilde{B} is sufficient fuzzy attribute of X .
- If $(X, \tilde{B}) \in \mathcal{G}_1 \cap \mathcal{G}_2$, then (X, \tilde{B}) is a sufficient and necessary fuzzy information granule, that is to say $X = \mathcal{P}(\tilde{B})$ and $\tilde{B} = \tilde{\mathcal{F}}(X)$ and \tilde{B} is a sufficient necessary fuzzy attribute of X .

- If $(X, \tilde{B}) \notin \mathcal{G}_1 \cup \mathcal{G}_2$, then (X, \tilde{B}) is an inconsistent information granule.

From Definition 1, we only consider the situation that there exist three fuzzy information granule spaces in $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$. Note that $\tilde{\mathcal{G}}_1 \cap \tilde{\mathcal{G}}_2$ be a fuzzy information granule space. Therefore, $\tilde{\mathcal{G}}_1 \cap \tilde{\mathcal{G}}_2$ is the concept space of $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$. However, $(a, \tilde{B}) \notin \tilde{\mathcal{G}}_1 \cup \tilde{\mathcal{G}}_2$ is not a fuzzy information granule of $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$. Moreover, if necessary, sufficient and necessary fuzzy information granules do not exist at the beginning of $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$. The approaches of volution of these fuzzy information granules are as follows.

Proposition 1 Let $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$ be a cognitive system, $\tilde{\mathcal{G}}_1$ be a necessary fuzzy information granule space, and $\tilde{\mathcal{G}}_2$ be a sufficient fuzzy information granule space. If $X \in L_1, \tilde{B} \in \tilde{L}_2$, then

- $(X \wedge \mathcal{P}(\tilde{B}), \tilde{B} \vee \tilde{\mathcal{F}}(X)) \in \tilde{\mathcal{G}}_1$
- $(X \vee \mathcal{P}(\tilde{B}), \tilde{B} \wedge \tilde{\mathcal{F}}(X)) \in \tilde{\mathcal{G}}_1$
- $(\mathcal{P}(\tilde{B}), \tilde{B} \wedge \tilde{\mathcal{F}}(X)) \in \tilde{\mathcal{G}}_1$
- $(X \wedge \mathcal{P}(\tilde{B}), \tilde{\mathcal{F}}(X)) \in \tilde{\mathcal{G}}_1$
- $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{B} \wedge \tilde{\mathcal{F}}(X)) \in \tilde{\mathcal{G}}_1$
- $(X \wedge \mathcal{P}(\tilde{B}), \tilde{\mathcal{F}}\mathcal{P}(\tilde{B})) \in \tilde{\mathcal{G}}_1$
- $(X \vee \mathcal{P}(\tilde{B}), \tilde{\mathcal{F}}\mathcal{P}(\tilde{B})) \in \tilde{\mathcal{G}}_2$
- $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{B} \vee \tilde{\mathcal{F}}(X)) \in \tilde{\mathcal{G}}_2$

Proposition 2 Let $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$ be a cognitive system, $\tilde{\mathcal{G}}_1$ be a necessary fuzzy information granule space, and $\tilde{\mathcal{G}}_2$ be a

Table 3 Attribute-oriented operator

$\{ \langle a, 1 \rangle, \langle b, 1 \rangle, \langle c, 1 \rangle, \langle d, 1 \rangle \}^* = \emptyset$	$\{ \langle a, 0.3 \rangle, \langle b, 0.3 \rangle, \langle c, 0.2 \rangle, \langle d, 0.5 \rangle \}^* = \{x_2, x_3\}$
$\{ \langle a, 0.6 \rangle, \langle b, 0.2 \rangle, \langle c, 0.5 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1\}$	$\{ \langle a, 0.2 \rangle, \langle b, 0.3 \rangle, \langle c, 0.4 \rangle, \langle d, 0.5 \rangle \}^* = \{x_2, x_4\}$
$\{ \langle a, 0.3 \rangle, \langle b, 0.3 \rangle, \langle c, 0.7 \rangle, \langle d, 0.5 \rangle \}^* = \{x_2\}$	$\{ \langle a, 0.2 \rangle, \langle b, 0.4 \rangle, \langle c, 0.2 \rangle, \langle d, 0.7 \rangle \}^* = \{x_3, x_4\}$
$\{ \langle a, 0.7 \rangle, \langle b, 0.6 \rangle, \langle c, 0.2 \rangle, \langle d, 0.9 \rangle \}^* = \{x_3\}$	$\{ \langle a, 0.3 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1, x_2, x_3\}$
$\{ \langle a, 0.2 \rangle, \langle b, 0.4 \rangle, \langle c, 0.4 \rangle, \langle d, 0.7 \rangle \}^* = \{x_4\}$	$\{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.4 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1, x_2, x_4\}$
$\{ \langle a, 0.3 \rangle, \langle b, 0.2 \rangle, \langle c, 0.5 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1, x_2\}$	$\{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1, x_2, x_3, x_4\}$
$\{ \langle a, 0.6 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1, x_3\}$	$\{ \langle a, 0.2 \rangle, \langle b, 0.3 \rangle, \langle c, 0.2 \rangle, \langle d, 0.5 \rangle \}^* = \{x_2, x_3, x_4\}$
$\{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.4 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1, x_2, x_4\}$	$\{ \langle a, 0.2 \rangle, \langle b, 0.2 \rangle, \langle c, 0.2 \rangle, \langle d, 0.3 \rangle \}^* = \{x_1, x_2, x_3, x_4\}$

sufficient fuzzy information granule space. If $(X_1, \tilde{B}_1) \in \tilde{\mathcal{G}}_1$ and $(X_2, \tilde{B}_2) \in \tilde{\mathcal{G}}_2$, then

1. $(X_1 \vee \mathcal{P}(\tilde{B}_1), \tilde{\mathcal{F}}(X_1 \vee \mathcal{P}(\tilde{B}_1))) \in \tilde{\mathcal{G}}_1 \cap \tilde{\mathcal{G}}_2$
2. $(\mathcal{P}(\tilde{B}_1 \vee \tilde{\mathcal{F}}(X_1)), \tilde{B}_1 \vee \tilde{\mathcal{F}}(X_1)) \in \tilde{\mathcal{G}}_1 \cap \tilde{\mathcal{G}}_2$
3. $(X_2 \wedge \mathcal{P}(\tilde{B}_2), \tilde{\mathcal{F}}(X_2 \wedge \mathcal{P}(\tilde{B}_2))) \in \tilde{\mathcal{G}}_1 \cap \tilde{\mathcal{G}}_2$
4. $(\mathcal{P}(\tilde{B}_2 \wedge \tilde{\mathcal{F}}(X_2)), \tilde{B}_2 \wedge \tilde{\mathcal{F}}(X_2)) \in \tilde{\mathcal{G}}_1 \cap \tilde{\mathcal{G}}_2$

Let $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$ be a cognitive system, $\tilde{\mathcal{G}}_1$ be a necessary fuzzy information granule space, and $\tilde{\mathcal{G}}_2$ be a sufficient fuzzy information granule space. If “ \vee ” and “ \wedge ” are defined operators of cognitive system, and

$$(X_1, \tilde{B}_1) \wedge (X_2, \tilde{B}_2) = (X_1 \wedge X_2, \tilde{\mathcal{F}}\mathcal{P}(\tilde{B}_1 \vee \tilde{B}_2))$$

$$(X_1, \tilde{B}_1) \vee (X_2, \tilde{B}_2) = (\mathcal{P}\tilde{\mathcal{F}}(X_1 \vee X_2), \tilde{B}_1 \wedge \tilde{B}_2)$$

Proposition 3 Let $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$ be a cognitive system, \mathcal{G}_1 be a necessary information granule space, \mathcal{G}_2 be a sufficient information granule space, and $\mathcal{G}_1 \cap \mathcal{G}_2$ be a sufficient and necessary information granule space. For arbitrary information granule $(X, \tilde{B}) \in \mathcal{G}_1 \cup \mathcal{G}_2$, there is only one sufficient and necessary information granule which is itself (i.e., (X, \tilde{B})); otherwise, we have two sufficient and necessary information granules, that is, $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{\mathcal{F}}(X))$ and $(\mathcal{P}(B), \tilde{\mathcal{F}}\mathcal{P}(B))$.

Proof Because $(L_1, \tilde{L}_2, \tilde{\mathcal{F}}, \mathcal{P})$ be a cognitive system, from Definition 1, we have three fuzzy information granule: \mathcal{G}_1 , \mathcal{G}_2 and $(\mathcal{G}_1 \cap \mathcal{G}_2)^c$, where $(\cdot)^c$ is the complement. Then, we divide it into three steps to prove it.

1. If $(X, \tilde{B}) \in \mathcal{G}_1$, from Definition 1, we have $X \leq \mathcal{P}(\tilde{B})$ and $\tilde{B} \leq L(X)$. Thus, $X \vee \mathcal{P}(\tilde{B}) = \mathcal{P}(\tilde{B})$, $\tilde{\mathcal{F}}(X \vee \mathcal{P}(\tilde{B})) = \tilde{\mathcal{F}}\mathcal{P}(\tilde{B})$, $\tilde{B} \vee \tilde{\mathcal{F}}(X) = \tilde{\mathcal{F}}(X)$, $\mathcal{P}(\tilde{B} \vee \tilde{\mathcal{F}}(X)) = \mathcal{P}\tilde{\mathcal{F}}(a)$; hence, two sufficient and necessary information granules are $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{\mathcal{F}}(X))$ and $(\mathcal{P}(\tilde{B}), \tilde{\mathcal{F}}\mathcal{P}(\tilde{B}))$.
2. If $(X, \tilde{B}) \in \mathcal{G}_2$, from Definition 1, we have $\tilde{\mathcal{F}}(X) \leq \tilde{B}$ and $\mathcal{P}(\tilde{B}) \leq X$. Thus, $X \wedge \mathcal{P}(\tilde{B}) = \mathcal{P}(\tilde{B})$, $\tilde{\mathcal{F}}(X \wedge \mathcal{P}(\tilde{B})) = \tilde{\mathcal{F}}\mathcal{P}(\tilde{B})$, $\tilde{B} \wedge \tilde{\mathcal{F}}(X) = \mathcal{P}(\tilde{B})$, $\mathcal{P}(\tilde{B} \wedge \tilde{\mathcal{F}}(X)) = \mathcal{P}\tilde{\mathcal{F}}(X)$; hence, two sufficient and necessary information granules are $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{\mathcal{F}}(X))$ and $(\mathcal{P}(\tilde{B}), \tilde{\mathcal{F}}\mathcal{P}(\tilde{B}))$.
3. If $(X, \tilde{B}) \notin \mathcal{G}_1 \cup \mathcal{G}_2$, it is immediate from Definition 1 and Proposition 2.

By combining 1), 2), and 3), this theorem is proven. #

Intuitively, Proposition 3 shows that $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{\mathcal{F}}(X))$ and $(\mathcal{P}(B), \tilde{\mathcal{F}}\mathcal{P}(B))$ is a sufficient and necessary information granule in the two-way learning system. $(X, \tilde{B}) = (\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{\mathcal{F}}(X)) = (\mathcal{P}(B), \tilde{\mathcal{F}}\mathcal{P}(B))$, if any $X \in L_1$ and $\tilde{B} \in L_2$, (X, \tilde{B}) is a sufficient and necessary information granule.

From the above discussion, we put forward a cognitive mechanism of transforming any information granule into sufficient and necessary fuzzy information granules as follows:

- **Case 1:** For any information granule (X, \tilde{B}) , its sufficient, necessary, and sufficient and necessary fuzzy information granules can be obtained with the two-way concept-cognitive learning method, and the detailed conversion process is provided in Propositions 1 and 2.
- **Case 2:** From the definition of the sufficient and necessary fuzzy information granule and cognitive operators, it is true that $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{\mathcal{F}}(X))$ is a fuzzy concept about a . So when one just knows the information from an object set X , a fuzzy concept can be obtained by computing $\mathcal{P}\tilde{\mathcal{F}}(X)$ and $\mathcal{P}\tilde{\mathcal{F}}(X)$.
- **Case 3:** From the definition of the sufficient and necessary fuzzy information granule and cognitive operators, it is obvious that $(\mathcal{P}(\tilde{B}), \tilde{\mathcal{F}}\mathcal{P}(\tilde{B}))$ is a fuzzy concept about \tilde{B} . So when one just knows the information from a fuzzy attribute set \tilde{B} , a fuzzy concept can be obtained by computing $\mathcal{P}(\tilde{B})$ and $\tilde{\mathcal{F}}\mathcal{P}(\tilde{B})$.

Concept Learning in Multi-source Fuzzy Decision Tables

It is well known that information about the same subject can be collected by different means with the development of technology. How to take full advantage of multi-source information to learn concepts is an imperative task during data processing. Considering the complexity of concept learning, we can divide the concept learning process of multi-source fuzzy decision tables into two steps. Firstly, multi-source information is fused based on some uncertainty measures. Secondly, fuzzy concepts are learned from the fused decision table.

Information from different sources is called multi-source information. In order to further deal with multi-source information, we first introduce a multi-source fuzzy decision table. The $MFI = \{\tilde{I}_{i \in N} | \tilde{I}_i = \{U, A_i, V_i, f_i\}\}$ can be called a multi-source fuzzy decision table with $|N|$ single fuzzy decision tables, where U is a finite non-empty set of object, A_i, V_i and f_i are the finite non-empty set of attributes, the domain of all objects under all attributes, and the information function in the i^{th} fuzzy decision table, respectively.

In each fuzzy information table $\tilde{I}_k = (U, A, V_k, f_k)$ ($k \in N$), for an arbitrary attribute $a \in A$, $\tilde{I}_k^a(x_i)$ ($k = 1, 2, \dots, s, i = 1, 2, \dots, n$) represents the value of x_i with respect to the attribute a in the k^{th} fuzzy decision table. And for any $x \in U$, a similarity class of x with respect to a in the \tilde{I}_k is defined as $T_a^k(x) = \{x_i | |\tilde{I}_k^a(x_i) - \tilde{I}_k^a(x)| \leq b\}$, where the threshold b is a given real number. It is obvious that $x_i \in T_a^k(x_i)$ and $x_i \in T_a^k(x_j) \Leftrightarrow x_j \in T_a^k(x_i)$.

For convenience, let $|N| = s$. $MFI = \{\tilde{I}_{i \in N} | \tilde{I}_i = \{U, A_i, V_i, f_i\}\}$ is composed of s single fuzzy decision table. In particular, if it is true that $A_i = A_j$ for all the $i \neq j$, then all the fuzzy decision tables have the same structure. Because the more information you have on the same thing learned knowledge should be more accurate. For the multi-source decision table, which has different structures, we may only need to find the ultimate goal as a middle bridge to establish the relationship between different information sources. However, the multi-source decision table, which has the same structure, will be required to achieve a higher goal in addition to the abovementioned middle bridge. So the research background of this paper is based on the multi-source fuzzy decision table which has the same structure and different information functions, namely, $MFI = \{\tilde{I}_{i \in N}^s | \tilde{I}_i = \{U, A, V_i, f_i\}\}$. For convenience, the multi-source decision table $MFI = \{\tilde{I}_{i \in N}^s | \tilde{I}_i = \{U, A, V_i, f_i\}\}$ is denoted as $MFI = \{\tilde{I}_1, \tilde{I}_2, \dots, \tilde{I}_s\}$. Unless otherwise specified, all multi-source fuzzy decision tables in this paper are defined as above shown, namely, $MFI = \{\tilde{I}_1, \tilde{I}_2, \dots, \tilde{I}_s\}$.

Multi-source Fuzzy Decision Fusion

By integrating data from different sources, the deficiency of single data can be made up. In order to make full use of multi-source information to realize the mutual complement and mutual confirmation of different information sources, all the signal fuzzy decision tables are fused based on the uncertainty measure. The detailed fusion process is shown in the following.

Conditional Entropy Fusion Mechanism

First of all, a critical issue is the reliability of information from different sources, namely the importance of information sources. In this paper, we select conditional entropy to measure the importance of information sources. A novel conditional entropy proposed by Dai et al. [26] is a great uncertainty measure that has monotonicity. The reason for selecting it is that the thinner the partition, the smaller the

conditional entropy. That is to say, with the decrease of conditional entropy, available knowledge may be increased. Therefore, the smaller the conditional entropy is, the more critical the source will be. In general, other uncertainty measures can also be used to determine the importance of information sources according to different decision goals.

According to the paper [26], this proposed conditional entropy can be used to evaluate the importance of attributes. Let $I = (U, A, V, f)$ be a decision table, where $U = \{x_1, x_2, \dots, x_n\}$ and $U/D = \{Y_1, Y_2, \dots, Y_m\}$. The conditional entropy (CE) of D regarding the attribute a in the I is defined as

$$H(D|a) = - \sum_{i=1}^n \sum_{j=1}^m \frac{|T_a(x_i) \cap Y_j|}{|U|} \log \frac{|T_a(x_i) \cap Y_j|}{|T_a(x_i)|}$$

where $T_a(x_i)$ is the similarity class of x_i with respect to a in the decision table I .

For the attribute a , the smaller the conditional entropy is, the more important the information source will be. So the following definitions in a fuzzy decision table are proposed.

Definition 2 Let $MFI = \{\tilde{I}_1, \tilde{I}_2, \dots, \tilde{I}_s\}$ be a multi-source fuzzy decision table, $U = \{x_1, x_2, \dots, x_n\}$ be the universe, and $U/D = \{Y_1, Y_2, \dots, Y_m\}$ be a classification of universe U . For any $a \in A$, the CE of D with respect to the attribute a in the \tilde{I}_k of MFI is defined as

$$H_a(D|\tilde{I}_k) = - \sum_{i=1}^{|U|} \sum_{j=1}^m \frac{|T_a^k(x_i) \cap Y_j|}{|U|} \log \frac{|T_a^k(x_i) \cap Y_j|}{|T_a^k(x_i)|}$$

where $T_a^k(x_i)$ is the similarity class of x_i with respect to a in the \tilde{I}_k .

The smaller the conditional entropy, the less uncertainty there is. Thus, we could select the best attribute among all multi-source according to the minimum principle and definition 2, as follows:

$$k_a = \arg \min_{k \in \{1, 2, \dots, s\}} (H_a(D|\tilde{I}_k))$$

Fig. 2 Multi-source fuzzy information fusion

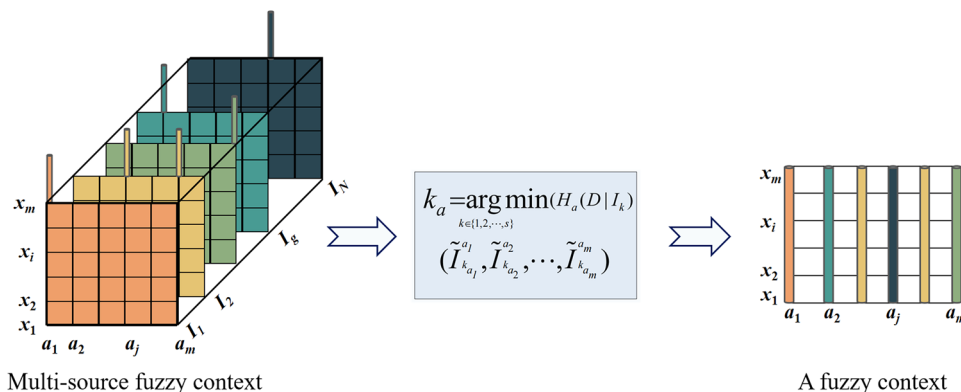


Table 4 Information source \tilde{I}_1

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	0.51	0.30	0.68	0.42	0.89	0.33
x_2	0.63	0.70	0.23	0.72	0.22	0.71
x_3	0.34	0.37	0.54	0.29	0.88	0.65
x_4	0.52	0.58	0.97	0.73	0.19	0.45
x_5	0.39	0.20	0.89	0.80	0.40	0.62
x_6	0.28	0.38	0.76	0.54	0.20	0.76
x_7	0.20	0.43	0.82	0.76	0.94	0.50
x_8	0.20	0.89	0.20	0.78	0.26	0.24
x_9	0.70	0.81	0.73	0.92	0.43	0.46
x_{10}	0.36	0.99	0.65	0.20	0.66	0.47

Then, the column of the attribute a in the k_a^{th} table is extracted as the column of the attribute a in the fused table. After this operation is performed on all the attributes, a new decision table can be obtained. Let $NI(a) = \tilde{I}_{k_a}^a, a \in A$ denote the operation on the attribute a and $NI = (\tilde{I}_{k_{a_1}}^{a_1}, \tilde{I}_{k_{a_2}}^{a_2}, \dots, \tilde{I}_{k_{a_m}}^{a_m})$ denote the new decision table, where m is the number of attribute.

The process of fusion is visually displayed in Fig. 2. There is a multi-source fuzzy decision table $MFI = \{\tilde{I}_1, \tilde{I}_2, \dots, \tilde{I}_s\}$ with s fuzzy decision table, and there are n objects and m attributes for each fuzzy decision table $\tilde{I}_i, (i = 1, 2, \dots, s)$. For each fuzzy decision table, we can calculate the conditional entropy of the decision attribute set with respect to each condition attribute by Definition 2. For example, there are s conditional entropy of the attribute a calculated in the MFI , namely, $H_a(D|\tilde{I}_1), H_a(D|\tilde{I}_2) \dots, H_a(D|\tilde{I}_s)$. According to Definition 2, the attribute a in the k_a^{th} table is extracted as the column of the attribute a in the fused table. In order to describe the results more vividly, we use different colors of rough lines to express the corresponding column extracted. These extracted columns construct a new fuzzy decision table. A more detailed presentation is shown in Fig. 2. Next, the proposed fusion theory is described by a specific example.

Example 1 This is a case study about the medical diagnosis. There are 10 patients who were suspected to be

infected with the *H1N1*. In order to diagnose, four hospitals examined the following six indicators, respectively. The results of testing in four hospitals are shown in Tables 4, 5, 6, and 7, where $A = \{a_1, a_2, \dots, a_6\}$ represent “White Blood Cell,” “Creatine Kinase,” “Aspartate Aminotransferase,” “Alanine Transaminase,” “Temperature,” and “Cough,” respectively. And $U = \{x_1, x_2, \dots, x_{10}\}$ denotes 10 those suspected patients. These tables obtain a multi-source fuzzy decision table, and their similarity classes can be get by Definition 2.

Given the partition of the universe is $UID = \{x_1, x_4, x_7\}, \{x_2, x_5, x_8\}, \{x_3, x_6, x_9, x_{10}\}$ and the value of b is 0.1, namely $b = 0.1$. The results of the conditional entropy of each attribute are shown in Table 8.

The smaller the conditional entropy, the more important the information source. According to the result of Table 8, the reliable source of attributes a_1, a_3, a_5 is all \tilde{I}_4 . Moreover, the reliable source of attributes a_2 and a_4 is all \tilde{I}_2 and the reliable source of attribute a_6 is \tilde{I}_3 . Then, a new fuzzy decision table can be established by extracting the corresponding column of each attribute. The detailed information of the new table is displayed in Table 9.

According to the theory of CE-fusion, an algorithm is designed for multi-source fuzzy information fusion. Detailed information is shown in Algorithm 1.

Table 5 Information source \tilde{I}_2

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	0.75	0.46	0.52	0.77	0.42	0.35
x_2	0.62	0.24	0.36	0.62	0.41	0.59
x_3	0.72	0.85	0.33	0.11	0.49	0.53
x_4	0.48	0.61	0.75	0.12	0.35	0.96
x_5	0.36	0.57	0.19	0.75	0.34	0.56
x_6	0.52	0.77	0.72	0.16	0.24	0.76
x_7	0.31	0.79	0.82	0.77	0.22	0.37
x_8	0.22	0.79	0.72	0.65	0.25	0.30
x_9	0.70	0.65	0.34	0.73	0.18	0.75
x_{10}	0.26	0.27	0.81	0.33	0.75	0.47

Algorithm 1: Algorithm for CE-fusion

```

Input : A multi-source fuzzy information system  $MFI = \{\tilde{I}_1, \tilde{I}_2, \dots, \tilde{I}_s\}$ , a classification  $U/D = \{Y_1, Y_2, \dots, Y_m\}$ ;
Output : A new information table.
1 begin
2   for  $k = 1 : s$  do
3     for each  $a \in A$  do
4       for each  $x \in U$  do
5         compute:  $T_a^k(x)$ ; // compute all  $T_a^k(x)$ ,
6         for any  $x \in U$  under attribute  $a$ ;
7       end
8        $HCE \leftarrow 0$ ;
9       for  $i = 1 : |U|$  do
10        for  $j = 1 : m$  do
11          if  $|T_a^k(x_i \cap Y_j)| > 0$  then
12             $HCE \leftarrow HCE - \frac{|T_a^k(x_i \cap Y_j)|}{|U|} \log \frac{|T_a^k(x_i \cap Y_j)|}{|T_a^k(x_i)|}$ ;
13          end
14        end
15      end
16       $H_a(D|\tilde{I}_k) \leftarrow HCE$ ; // record CE under attribute  $a$  with
17      the  $k^{th}$  information source;
18    end
19    for each  $a \in A$  do
20       $minCE \leftarrow +\infty$ ;
21      for  $k = 1 : s$  do
22        if  $H_a(D|\tilde{I}_k) < minCE$  then
23           $minCE \leftarrow H_a(D|\tilde{I}_k)$ ;
24           $k_a \leftarrow k$ ;
25        end
26      end
27    return:  $(\tilde{I}_{k_{a_1}}^{a_1}, \tilde{I}_{k_{a_2}}^{a_2}, \dots, \tilde{I}_{k_{a_{|A|}}}^{a_{|A|}})$ .
28 end

```

In Algorithm 1, we first calculate all the similarity class $T_a^k(x)$ of any $x \in U$ under attribute a in the k^{th} fuzzy decision table. Then, the conditional entropy $H_a(D|\tilde{I}_k)$ be calculated in the k^{th} fuzzy decision table. Finally, the fuzzy decision table in which the conditional entropy of D with respect to a is minimal is selected as the reliable source of the attribute a . Then, a new fuzzy decision table can be established by extracting the corresponding column of each attribute, namely $NI = (\tilde{I}_{k_{a_1}}^{a_1}, \tilde{I}_{k_{a_2}}^{a_2}, \dots, \tilde{I}_{k_{a_{|A|}}}^{a_{|A|}})$. Taking into account the efficiency of the algorithm, we analyze the complexity of Algorithm 1, which is shown in Table 10.

In Algorithm 1, we compute all $T_a^k(x)$, for any $x \in U$ under attribute a in steps 4–5. Steps 6–14 calculate the

conditional entropy for any attribute $a \in A$ in the k^{th} information source. Steps 17–26 find the minimum conditional entropy of each attribute as their reliable sources. At last, return the results.

Uncertainty Measure of CE-Fusion

In order to show the advantages of the CE-fusion method, approximate precision and approximate quality are introduced to measure our methods. First, the definition of approximate precision is proposed in the following.

Table 6 Information source \tilde{I}_3

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	0.66	0.75	0.17	0.43	0.62	0.58
x_2	0.71	0.65	0.99	0.73	0.40	0.85
x_3	0.70	0.43	0.58	0.14	0.46	0.51
x_4	0.77	0.10	0.88	0.73	0.56	0.29
x_5	0.43	0.44	0.30	0.79	0.41	0.65
x_6	0.22	0.27	0.36	0.59	0.31	0.84
x_7	0.20	0.72	0.81	0.33	0.30	0.44
x_8	0.09	0.46	0.27	0.63	0.26	0.37
x_9	0.70	0.54	0.44	0.57	0.34	0.56
x_{10}	0.17	0.84	0.53	0.76	0.85	0.47

Table 7 Information source \tilde{I}_4

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	0.06	0.69	0.37	0.69	0.23	0.49
x_2	0.72	0.38	0.34	0.20	0.89	0.55
x_3	0.28	0.35	0.78	0.10	0.69	0.66
x_4	0.54	0.12	0.60	0.59	0.92	0.48
x_5	0.45	0.78	0.76	0.20	0.75	0.73
x_6	0.15	0.68	0.37	0.96	0.20	0.85
x_7	0.34	0.88	0.23	0.80	0.33	0.48
x_8	0.27	0.58	0.89	0.13	0.26	0.31
x_9	0.58	0.55	0.89	0.20	0.99	0.32
x_{10}	0.43	0.74	0.39	0.65	0.67	0.56

Table 8 Conditional entropy of each attribute in the *MF1*

\tilde{I}	a_1	a_2	a_3	a_4	a_5	a_6
\tilde{I}_1	0.6502	0.5257	0.8476	1.0549	0.5885	0.8553
\tilde{I}_2	0.5649	0.3452	1.0095	0.9303	1.1399	0.8318
\tilde{I}_3	0.8808	0.5493	0.7337	0.9332	1.1747	0.7487
\tilde{I}_4	0.0863	0.4122	0.3452	1.1596	0.4970	0.9020

Definition 3 Let $\tilde{I} = (U, AT)$ be a fuzzy information system, where $U/D = \{Y_1, Y_2, \dots, Y_m\}$, $\tilde{I}^a(x_j) (j = 1, 2, \dots, n)$ is the value of x_j with respect to a in the \tilde{I} . Given the constant θ , for any $Y_i \in U/D (j = 1, 2, \dots, m)$, the lower and upper approximations of Y_j are respectively defined as

$$\underline{Apr}(Y_i) = \{x \in U | T_\theta(x) \subseteq Y_i\},$$

$$\overline{Apr}(Y_i) = \{x \in U | T_\theta(x) \cap Y_i \neq \emptyset\}$$

where $T_\theta(x) = \{x_i | \sum_{a \in AT} |\tilde{I}^a(x_i) - \tilde{I}^a(x)| \leq \theta\}$ is the similarity class of x with respect to a in the \tilde{I} . The lower approximation and upper approximation of U/D are $\underline{Apr}(U/D) = \underline{Apr}(Y_1) \cup \underline{Apr}(Y_2) \cup \dots \cup \underline{Apr}(Y_m)$ and $\overline{Apr}(U/D) = \overline{Apr}(Y_1) \cup \overline{Apr}(Y_2) \cup \dots \cup \overline{Apr}(Y_m)$, respectively.

The approximation precision (AP) and approximation quality (AQ) of U/D and in the \tilde{I} can be defined as follows:

$$AP_R(U/D) = \frac{\sum_{Y_i \in U/D} |\underline{Apr}(Y_i)|}{\sum_{Y_i \in U/D} |\overline{Apr}(Y_i)|}$$

$$AQ_R(U/D) = \frac{\sum_{Y_i \in U/D} |\underline{Apr}(Y_i)|}{|U|}$$

Then, the above theory is expounded by specific examples. Followed by example 1 and example 2, the approximation precision of U/D in the new tables of CE-fusion is calculated. Due to different science fields may require different similarity class thresholds, in this paper, the similarity class

Table 9 The new decision table after CE-fusion

U	a_1	a_2	a_3	a_4	a_5	a_6
x_1	0.06	0.46	0.37	0.77	0.23	0.58
x_2	0.72	0.24	0.34	0.62	0.89	0.85
x_3	0.28	0.85	0.78	0.11	0.69	0.51
x_4	0.54	0.61	0.60	0.12	0.92	0.29
x_5	0.45	0.57	0.76	0.75	0.75	0.65
x_6	0.15	0.77	0.37	0.16	0.20	0.84
x_7	0.34	0.79	0.23	0.77	0.33	0.44
x_8	0.27	0.79	0.89	0.65	0.26	0.37
x_9	0.58	0.65	0.89	0.73	0.99	0.56
x_{10}	0.43	0.27	0.39	0.33	0.67	0.47

Table 10 The computational complexity of Algorithm 1

Steps 4–5	$O(U ^2)$
Steps 6–14	$O(U \times m^2)$
Steps 1–16	$O(s \times A \times (U ^2 + U \times m^2))$
Steps 17–25	$O(A \times s)$
Step 26	$O(U \times A)$
Total	$O(s \times A \times (U ^2 + U \times m^2) + A \times s + U \times A)$

threshold θ varies from 0.3 to 1.4 with an increase of 0.1. The experimental results of CE-fusion are shown in Table 11.

Two-way Concept-Cognitive Learning of *MF1*

Concept learning is an important part of human cognition. People from different science fields have different means of information cognition. There are different approaches to learning concepts. Let (U, A, \tilde{I}) be a fuzzy formal context, $X \subseteq U$ be an object set, and \tilde{B} be a fuzzy attribute set. There are 3 cases about concept learning from different perspectives.

- **Case 1’:** For any information granule (X, \tilde{B}) , its sufficient, necessary, and sufficient and necessary fuzzy information granules can be obtained with the two-way concept-cognitive learning method, and the detailed conversion process is provided in the “Two-way Concept-Cognitive Learning” section.
- **Case 2’:** From the definition of the fuzzy concept and the property (3) of cognitive operators, it is true that $(\mathcal{P}\tilde{\mathcal{F}}(X), \tilde{\mathcal{F}}(X))$ is a fuzzy concept about X . So when one just knows the information from an object set X , a fuzzy concept can be obtained by computing $\mathcal{P}\tilde{\mathcal{F}}(X)$ and $\tilde{\mathcal{F}}(X)$. The corresponding calculation process is realized by Algorithm 2.
- **Case 3’:** According to the definition of the fuzzy concept and the property (3) of cognitive operators, it is obvious that $(\mathcal{P}(B), \tilde{\mathcal{F}}\mathcal{P}(B))$ is a fuzzy concept about \tilde{B} . So when one just knows the information from a fuzzy attribute set \tilde{B} , a fuzzy concept can be obtained by computing $\mathcal{P}(B)$ and $\tilde{\mathcal{F}}\mathcal{P}(B)$. The corresponding calculation process is realized by Algorithm 3.

When we only know the extent of a concept, it is our primary task to learn the intent of concept. The process of object-oriented concept learning is realized by Algorithm 2.

In Algorithm 2, step 2 initializes the intent and extent of the concept to be \emptyset , steps 5–7 seek attribute of the intent of the concept which we will learn, step 8 records the intent of the target concept to B , steps 11–17 search object of the extent of the target concept, and steps 18–20 add the object that is searched in steps 11–17 to Y . At last, return the result.

Table 11 AP and AQ of CE-fusion method

θ	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4
AP	1.0000	1.0000	1.0000	1.0000	1.0000	0.6667	0.6667	0.6667	0.4286	0.4286	0.2500	0.0000
AQ	1.0000	1.0000	1.0000	1.0000	1.0000	0.8000	0.8000	0.8000	0.6000	0.6000	0.4000	0.0000

Algorithm 2: Fuzzy concept learning oriented by an object set

```

Input : A fuzzy information system  $\tilde{I}S = \{U, A, \tilde{I}\}$  and an object set  $X \subseteq U$ ;
Output : A fuzzy concept.
1 begin
2    $B \leftarrow \phi, Y \leftarrow \phi$ ;
3   for each  $a \in A$  do
4      $\mu_{\tilde{X}}(a) \leftarrow 1$ ;
5     for each  $x \in X$  do
6        $\mu_{\tilde{X}}(a) \leftarrow \mu_{\tilde{X}}(a) \wedge \mu_{\tilde{I}}(x, a)$ ;           // seek intent of target concept ;
7     end
8      $B \leftarrow B \cup \{ \langle a, \mu_{\tilde{X}}(a) \rangle \}$ ;           // record intent of target concept ;
9   end
10  for each  $x \in U$  do
11     $flag \leftarrow \text{true}$ ;
12    for each  $a \in A$  do
13      if  $\mu_{\tilde{I}}(x, a) < B(a)$  then
14         $flag \leftarrow \text{false}$ ;           // seek extent of target concept ;
15        break;
16      end
17    end
18    if  $flag$  then
19       $Y \leftarrow Y \cup \{x\}$ ;           // record intent of target concept ;
20    end
21  end
22  return:  $(Y, B)$ .
23 end

```

Moreover, the computational complexity of Algorithm 2 is analyzed, which is shown in Table 12.

When we only know the intent of a concept, it is our main task to learn the extent of the concept. The process

of attribute-oriented concept learning is realized by Algorithm 3. By Algorithm 3, we know how to learn a fuzzy concept by an attribute set.

Algorithm 3: Fuzzy concept learning oriented by a fuzzy attribute set.

```

Input : A fuzzy information system  $\tilde{I}S = \{U, A, \tilde{I}\}$  and a fuzzy attribute set  $A \subseteq AT$ ;
Output : A fuzzy concept.
1 begin
2    $B \leftarrow \phi, Y \leftarrow \phi$ ;
3   for each  $x \in U$  do
4      $flag \leftarrow \text{true}$ ;
5     for each  $a \in A$  do
6       if  $\mu_{\tilde{I}}(x, a) < A(a)$  then
7          $flag \leftarrow \text{false}$ ;           // seek extent of target concept ;
8         break;
9       end
10    end
11    if  $flag$  then
12       $Y \leftarrow Y \cup \{x\}$ ;           // record extent of target concept ;
13    end
14  end
15  for each  $a \in A$  do
16     $\mu_{\tilde{Y}}(a) \leftarrow 1$ ;
17    for each  $x \in Y$  do
18       $\mu_{\tilde{Y}}(a) \leftarrow \mu_{\tilde{Y}}(a) \wedge \mu_{\tilde{I}}(x, a)$ ;           // seek intent of target concept ;
19    end
20     $B \leftarrow B \cup \{ \langle a, \mu_{\tilde{Y}}(a) \rangle \}$ ;           // record intent of target concept ;
21  end
22  return:  $(Y, B)$ .
23 end

```

Table 12 The computational complexity of Algorithm 2

Steps 1–9	$O(AT \times X)$
Steps 10–21	$O(U \times AT)$
Total	$O(AT \times X + U \times AT)$

According to the discussion of Algorithm 3, we also analyze its complexity, which is shown in Table 13.

From the computation steps and the computational complexity, Algorithm 3 is similar to Algorithm 2. Step 2 initializes intent and extent of the concept, which we will learn to be ϕ , steps 4–10 search the object of the extent of the target concept, and steps 11–13 record the object of the extent to Y . Steps 16–19 seek attribute of the intent of the target concept, and step 20 records the attribute of the intent to B . At last, return the result.

Currently, there is no unified standard in the multi-source fuzzy decision table. Based on the uncertainty measure in the rough set, we defined AP and AQ of concept as follows.

Definition 4 Let (U, A, \tilde{I}) be a fuzzy formal context, and (X, \tilde{B}) be a fuzzy concept. The AP and AQ of extent of (X, \tilde{B}) are defined as:

$$AP_{T_\theta}(X, \tilde{B}) = \frac{|Apr(X, \tilde{B})|}{|\overline{Apr}(X, \tilde{B})|},$$

$$AQ_{T_\theta}(X, \tilde{B}) = \frac{|Apr(X, \tilde{B})|}{|U|}$$

where $Apr(X, \tilde{B}) = \{x \in U | T_\theta(x) \subseteq X\}$ and $\overline{Apr}(X, \tilde{B}) = \{x \in U | T_\theta(x) \cap X \neq \phi\}$ are the lower and upper approximations of extent X of concept (X, \tilde{B}) and $T_\theta(x) = \{x_i | \sum_{a \in AT} |\tilde{I}^a(x_i) - \tilde{I}^a(x)| \leq \theta\}$ is the similarity class of x about a in the \tilde{I} .

In order to display the advantage of concept learning based on CE-fusion in a multi-source fuzzy decision table, the AP and AQ of the extent of concepts are compared in a specific instance. The detailed process is shown in the following.

Example 2 According to the result of CE-fusion shown in Table 9 and the method of concept learning described in the “Two-way Concept-Cognitive Learning” section, the concepts related to the decision classes can be learned, which are

Table 13 The computational complexity of Algorithm 3

Steps 1–14	$O(U \times AT)$
Steps 15–21	$O(AT \times A^*)$
Total	$O(U \times AT + AT \times A^*)$

$$(X_1, \tilde{A}_1) \triangleq (\{x_1, x_4, x_7\}, \{\langle a_1, 0.19 \rangle, \langle a_2, 0.26 \rangle, \langle a_3, 0.71 \rangle, \langle a_4, 0.20 \rangle, \langle a_5, 0.46 \rangle, \langle a_6, 0.08 \rangle\}),$$

$$(X_2, \tilde{A}_2) \triangleq (\{x_2, x_5, x_8\}, \{\langle a_1, 0.13 \rangle, \langle a_2, 0.19 \rangle, \langle a_3, 0.20 \rangle, \langle a_4, 0.59 \rangle, \langle a_5, 0.26 \rangle, \langle a_6, 0.37 \rangle\}),$$

$$(X_3, \tilde{A}_3) \triangleq (\{x_3, x_6, x_9, x_{10}\}, \{\langle a_1, 0.28 \rangle, \langle a_2, 0.34 \rangle, \langle a_3, 0.16 \rangle, \langle a_4, 0.18 \rangle, \langle a_5, 0.20 \rangle, \langle a_6, 0.47 \rangle\}).$$

Given $\theta = 1.4$, the lower and upper approximations of extents of above concepts can be calculated as follows:

$$\underline{Apr}(X_1, \tilde{A}_1) = \{x_4, x_7\}, \overline{Apr}(X_1, \tilde{A}_1) = \{x_1, x_3, x_4, x_5, x_7\},$$

$$\underline{Apr}(X_2, \tilde{A}_2) = \{x_8, x_9\}, \overline{Apr}(X_2, \tilde{A}_2) = \{x_1, x_2, x_5, x_6, x_8, x_9\},$$

$$\underline{Apr}(X_3, \tilde{A}_3) = \{x_9, x_{10}\}, \overline{Apr}(X_3, \tilde{A}_3) = \{x_1, x_2, x_3, x_6, x_9, x_{10}\}.$$

Then according to the above discussion, the approximation precision of extents of the above concepts can be obtained, namely $AP_{T_\theta}(X_1, \tilde{A}_1) = 2/5$, $AP_{T_\theta}(X_2, \tilde{A}_2) = 1/3$, $AP_{T_\theta}(X_3, \tilde{A}_3) = 1/3$; $AQ_{T_\theta}(X_1, \tilde{A}_1) = 1/5$, $AQ_{T_\theta}(X_2, \tilde{A}_2) = 1/5$, $AQ_{T_\theta}(X_3, \tilde{A}_3) = 1/5$.

According to the results of concept learning, we can find that concepts are different for the same decision class in the new tables of CE-fusion. From the perspective of the approximation precision of the extent of concepts, CE-fusion in the multi-source fuzzy decision table is, to some extent, a suitable fusion method. Therefore, CE-fusion can be used as an effective technique in multiple cognition. Concept learning based on CE-fusion is a suitable concept learning method in the multi-source fuzzy decision table.

Experiment Evaluations

In this section, we first introduce the experimental setting and multi-source fuzzy context in the “Experimental Settings” subsection. In the “Outcome Evolution” subsection, we evaluate the outcome of the presented CE-fusion way for concept learning. Finally, we also reveal the effectiveness of CE-fusion with other methods on the public dataset.

Experimental Settings

In order to evaluate the effectiveness of our method for concept-cognitive learning in a multi-source fuzzy context, we conduct a series of experiments on public datasets from Keel and UCI Repository, namely, “User Knowledge Modeling,” “Balance,” “Pima,” “Vehicle,” “Winequality_red,”

Table 14 The basic information of multi-source fuzzy context

ID	Dataset name	Instance	Feature	Class
1	User Knowledge Modeling	4030	6	4
2	Balance	6250	4	3
3	Pima	7680	8	2
4	Vehicle	8460	18	4
5	Winequality_red	15,990	11	11
6	Wifi_localization_ok Objects	20,000	7	4

and “Wifi_localization_ok Objects.” Note that ten sources for each dataset are generated by blurring the original dataset and adding random noise to the original dataset. Detailed information about multi-source fuzzy context is shown in Table 14. This experimental program runs on a personal computer with hardware and software parameters, as shown in Table 15.

The method of generating multi-source fuzzy decision context is proposed in the following. Firstly, to obtain a fuzzy decision context, each column of data is divided by the maximum value of the column in the original context. Then, a multi-source fuzzy decision context is constructed by adding Gauss and random noise to the fuzzy decision context.

Let $MFI = \{\tilde{I}_1, \tilde{I}_2, \dots, \tilde{I}_s\}$ be a multi-source fuzzy decision context constructed by a fuzzy decision context \tilde{I} . Firstly, s numbers (g_1, g_2, \dots, g_s) which obey the $N(0, \sigma)$ distribution are generated, where σ is the standard deviation. The method of adding Gauss noise is defined as follows:

$$\tilde{I}_i(x, a) = \begin{cases} \tilde{I}(x, a) + g_i & \text{if } (0 \leq \tilde{I}(x, a) + g_i \leq 1) \\ \tilde{I}(x, a) & \text{else} \end{cases}$$

where $\tilde{I}(x, a)$ is the value of object x under attribute a in fuzzy decision context, and $\tilde{I}_i(x, a)$ is the value of object x under attribute a in the i^{th} fuzzy decision context \tilde{I}_i .

Then, s random numbers (e_1, e_2, \dots, e_s) are generated and these numbers are between $-e$ and e , where e is random error threshold. The method of adding random noise is given in the following.

$$\tilde{I}_i(x, a) = \begin{cases} \tilde{I}(x, a) + e_i & \text{if } (0 \leq \tilde{I}(x, a) + e_i \leq 1) \\ \tilde{I}(x, a) & \text{else} \end{cases}$$

Table 15 The description of experiment environment

Name	Model	Parameters
CPU	Intel i7 – 10750H	2.60GHz
Memory	Lenovo DDR4	32GB
Hard disk	Lexar Data	1TB
System	Windows 10	64 bit
Platform	Python	3.7

where $\tilde{I}(x, a)$ represents value of object x under attribute a in fuzzy decision table, and $\tilde{I}_i(x, a)$ represents object x under attribute a in the i^{th} fuzzy information source \tilde{I}_i .

Next, 35% objects are randomly selected from the fuzzy decision context \tilde{I} , and then Gauss noise is added to these objects. Thirty percent of objects are randomly selected from the rest of the context, and random noise is added. Finally, $MFI = \{\tilde{I}_1, \tilde{I}_2, \dots, \tilde{I}_s\}$ can be obtained.

Outcome Evolution

In order to test the validity of fuzzy decision context after the CE-fusion method and concept learning mechanism, we select nine fuzzy set membership functions to conduct a series of experiments. Gauss form, Cauchy form, and Γ form membership function are selected as the representation of fuzzy membership functions, which are divided into small, middle, and large types. In this process, drop half Cauchy form, Cauchy form, and L half Cauchy form are small type, middle, and large type, respectively. These nine different membership functions considered in this paper are small, middle, and large fuzzy membership functions of Gaussian, Cauchy, and Γ . For convenience, symbols SG , MG , LG , SC , MC , LC , $S\Gamma$, $M\Gamma$, and $L\Gamma$ denote the nine membership functions. Corresponding graphics are shown in Fig. 3.

Based on the fuzzy membership functions and Algorithm 3, evaluate our methods' approximation precision and approximation quality on six datasets (i.e., datasets 1–6). The details are shown in Tables 16, 17, 18, 19, 20, and 21. In different science fields, the standard deviation of Gauss noise and random error threshold of random noise may be different. In this paper, experiments are carried out five times for each dataset and let the standard deviation σ and random error threshold e vary from 0.005 to 0.03 with an increase of 0.005 every time, denoted as noises 1–5. In this experiment, the thresholds b and θ are (0.1, 0.4).

Tables 16, 17, 18, 19, 20, and 21 record the measures values of five noises under nine membership functions, where AP and AQ represent the value of approximation precision and approximation quality. From these tables, we could find that all the measures (approximation precision and approximation quality) are higher than those of M-fusion methods on different datasets. All in all, based on the concept extension's approximate precision and approximate precision, the conditional entropy fusion method (CE-fusion) and the mean value fusion method are compared in the experiment. The results show that the CE-fusion method is good in the concept learning of multi-source fuzzy decision tables.

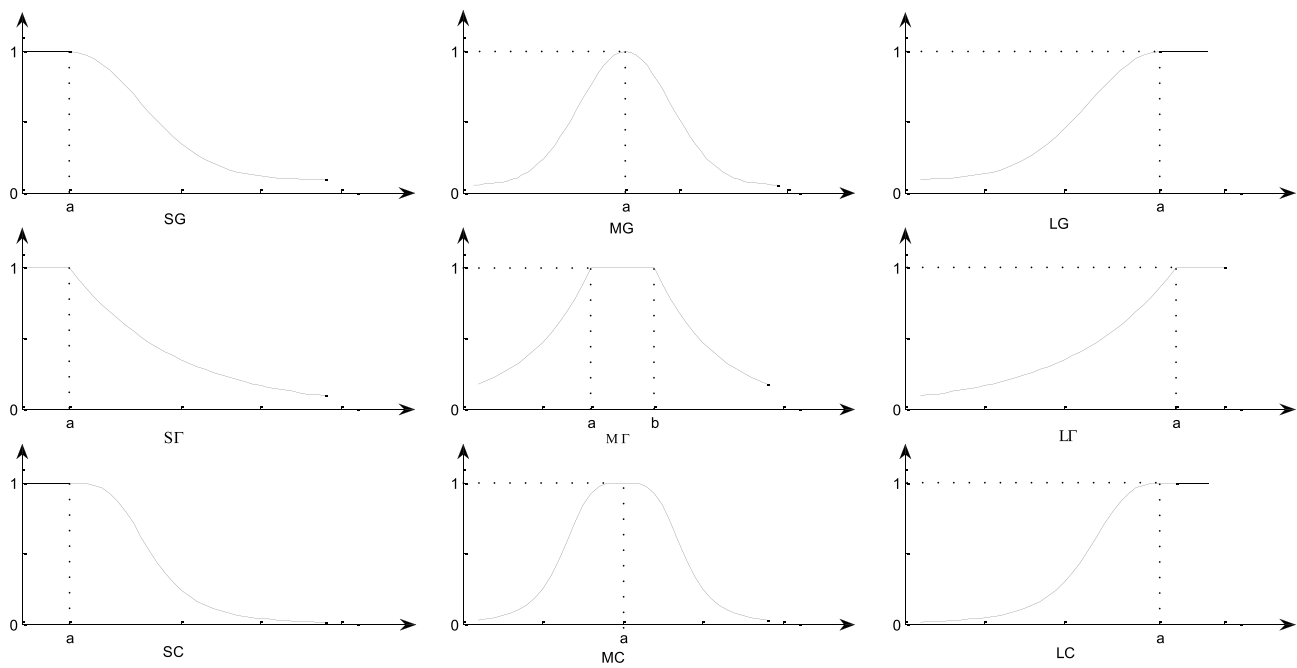


Fig. 3 The nine kinds of membership functions

Parameters Analysis

In order to verify the effectiveness of the CE-Fusion method, in this subsection, we also compare and record the approximation precision and approximation quality of four fusion methods under different noises. Figures 4, 5, 6, 7, 8, and 9 show the approximation accuracy of four fusion methods, where the dark blue, light blue, orange, and red lines represent the approximation accuracy on six datasets under different noises of M-fusion, min-fusion, max-fusion, and CE-fusion. According to these pictures, we find that in most noises cases, the height of the yellow bar is higher than or equal to the other three bars, and the red line is higher than those of other colors, which illustrates the effectiveness of the CE-fusion method than the other three compared methods under different noises.

In addition, we also take the approximation quality to analyze the advantage of the proposed method under different noises. The experimental results of approximation quality on six datasets are shown in Figs. 10, 11, 12, 13, 14, and 15 under nine membership functions. The graphic parameters are the same as those above. In these figures, we could find that the approximation quality is mostly higher than that of the other three compared methods except some cases. For Balance and Wifi_localization_ok Objects, the approximation quality of CE-fusion is all higher than that of the other three compared methods. Meanwhile, for User Knowledge Modeling, Pima, Vehicle, and Winequality_red, respectively, the CE-fusion method is higher than other M-fusions for 6, 6, 8, and 7 times in 9 experiments. All the experimental results verify the superiority of the CE-fusion method under different noises.

Table 16 The average value of AP and AQ on User Knowledge Modeling Dataset

Method	M-fusion		Min-fusion		Max-fusion		CE-fusion	
	AP	AQ	AP	AQ	AP	AQ	AP	AQ
SG	0.4946	0.1395	0.4570	0.1116	0.4967	0.1589	0.6808	0.1628
MG	0.5265	0.3620	0.5376	0.3395	0.5505	0.4008	0.7551	0.4465
LG	0.4008	0.1674	0.3833	0.1349	0.4700	0.2349	0.6520	0.2349
SC	0.3650	0.0891	0.2494	0.0512	0.4243	0.1186	0.5579	0.1124
MC	0.4782	0.3147	0.4756	0.2891	0.5217	0.3667	0.7332	0.4031
LC	0.7064	0.1062	0.8998	0.0961	0.6272	0.1093	0.9147	0.1124
SΓ	0.4338	0.1194	0.4041	0.0938	0.4889	0.1566	0.6674	0.1527
MΓ	0.5104	0.3620	0.5150	0.3395	0.5316	0.4008	0.7315	0.4465
LΓ	0.4766	0.2450	0.4100	0.1907	0.4932	0.2977	0.6992	0.3225

Table 17 The average value of AP and AQ on Balance Dataset

Method	M-fusion		Min-fusion		Max-fusion		CE-fusion	
	AP	AQ	AP	AQ	AP	AQ	AP	AQ
<i>SG</i>	0.2151	0.1536	0.3543	0.2083	0.3719	0.2480	0.5552	0.2915
<i>MG</i>	0.2650	0.2442	0.3696	0.2826	0.4371	0.3600	0.5747	0.3994
<i>LG</i>	0.2411	0.2160	0.3592	0.2749	0.3821	0.2925	0.5745	0.3754
<i>SC</i>	0.1350	0.0864	0.2706	0.1405	0.2522	0.1453	0.4542	0.2003
<i>MC</i>	0.1582	0.1296	0.2887	0.1923	0.3090	0.2256	0.4740	0.2730
<i>LC</i>	0.2411	0.2160	0.3594	0.2806	0.3828	0.2931	0.5720	0.3757
<i>ST</i>	0.1350	0.0864	0.2706	0.1405	0.2707	0.1616	0.4617	0.2058
<i>MT</i>	0.1582	0.1299	0.2877	0.1923	0.3362	0.2544	0.4787	0.2829
<i>LT</i>	0.2411	0.2160	0.3592	0.2790	0.3821	0.2925	0.5725	0.3754

Table 18 The average value of AP and AQ on Pima Dataset

Method	M-fusion		Min-fusion		Max-fusion		CE-fusion	
	AP	AQ	AP	AQ	AP	AQ	AP	AQ
<i>SG</i>	0.6184	0.3469	0.6174	0.3315	0.6468	0.3768	0.7356	0.3758
<i>MG</i>	0.6184	0.3469	0.6174	0.3315	0.6468	0.3768	0.7356	0.3758
<i>LG</i>	0.8986	0.0276	0.8575	0.0201	0.8637	0.0297	1.0000	0.0393
<i>SC</i>	0.5208	0.1341	0.5065	0.1250	0.4854	0.1495	0.6350	0.2247
<i>MC</i>	0.5208	0.1341	0.5065	0.1250	0.4854	0.1495	0.6350	0.2247
<i>LC</i>	0.8710	0.0352	0.8594	0.0253	0.8182	0.0466	0.9640	0.0346
<i>ST</i>	0.4998	0.1635	0.4895	0.1289	0.4760	0.1625	0.6050	0.2445
<i>ST</i>	0.4998	0.1635	0.4871	0.1289	0.4665	0.1625	0.6050	0.2469
<i>ST</i>	0.8786	0.0300	0.8290	0.0203	0.8650	0.0398	1.0000	0.0305

Table 19 The average value of AP and AQ on Vehicle dataset

Method	M-fusion		Min-fusion		Max-fusion		CE-fusion	
	AP	AQ	AP	AQ	AP	AQ	AP	AQ
<i>SG</i>	0.9898	0.7539	0.9936	0.6955	0.9881	0.7875	0.9965	0.7475
<i>MG</i>	0.9878	0.7487	0.9936	0.6955	0.9881	0.7442	0.9965	0.7870
<i>LG</i>	0.9831	0.1929	0.9970	0.1187	0.9847	0.1979	0.9977	0.3296
<i>SC</i>	0.9868	0.6570	0.9946	0.5667	0.9870	0.6381	0.9959	0.7201
<i>MC</i>	0.9868	0.6570	0.9946	0.5667	0.9870	0.6381	0.9959	0.7201
<i>LC</i>	0.9714	0.2087	0.9576	0.1286	0.9705	0.2149	0.9881	0.3444
<i>ST</i>	0.9897	0.7504	0.9936	0.6863	0.9881	0.7447	0.9965	0.7863
<i>ST</i>	0.9897	0.7511	0.9936	0.6891	0.9879	0.7896	0.9965	0.7485
<i>ST</i>	0.9747	0.1995	0.9745	0.1236	0.9765	0.2050	0.9977	0.3345

Table 20 The average value of AP and AQ on Winequality_red dataset

Method	M-fusion		Min-fusion		Max-fusion		CE-fusion	
	AP	AQ	AP	AQ	AP	AQ	AP	AQ
<i>SG</i>	0.7672	0.6519	0.6999	0.5611	0.8298	0.7001	0.8568	0.7406
<i>MG</i>	0.7671	0.6519	0.6977	0.5582	0.8290	0.6987	0.8557	0.7398
<i>LG</i>	0.3719	0.0949	0.3454	0.0505	0.4742	0.1152	0.6045	0.2059
<i>SC</i>	0.6965	0.5548	0.5909	0.4148	0.7653	0.6494	0.8311	0.6029
<i>MC</i>	0.6965	0.5548	0.5909	0.4148	0.7653	0.6029	0.8311	0.6494
<i>LC</i>	0.4077	0.5548	0.3601	0.4148	0.4889	0.6494	0.6204	0.6029
<i>ST</i>	0.7526	0.6231	0.6784	0.5221	0.7919	0.6655	0.8526	0.6944
<i>MT</i>	0.7589	0.6324	0.6805	0.5246	0.7918	0.6690	0.8533	0.6974
<i>LT</i>	0.3930	0.1107	0.3438	0.0537	0.4813	0.1253	0.5997	0.2221

Table 21 The average value of AP and AQ on Wifi_localization_ok Objects dataset

Method	M-fusion		Min-fusion		Max-fusion		CE-fusion	
	AP	AQ	AP	AQ	AP	AQ	AP	AQ
<i>SG</i>	0.5439	0.5219	0.5325	0.5009	0.5666	0.5539	0.6250	0.5740
<i>MG</i>	0.5439	0.5219	0.5325	0.5009	0.5666	0.5539	0.6250	0.5740
<i>LG</i>	0.4809	0.4579	0.4389	0.4099	0.5353	0.5176	0.5904	0.5356
<i>SC</i>	0.5446	0.5215	0.5325	0.5009	0.5665	0.5536	0.6251	0.5739
<i>MC</i>	0.5446	0.5215	0.5325	0.5009	0.5665	0.5536	0.6251	0.5739
<i>LC</i>	0.5212	0.5067	0.4672	0.4437	0.5780	0.5652	0.6212	0.5797
$S\Gamma$	0.5437	0.5215	0.5325	0.5009	0.5665	0.5537	0.6250	0.5740
$M\Gamma$	0.5437	0.5222	0.5335	0.5023	0.5690	0.5565	0.6311	0.5805
$L\Gamma$	0.4894	0.4698	0.4495	0.4242	0.5566	0.5403	0.6041	0.5574

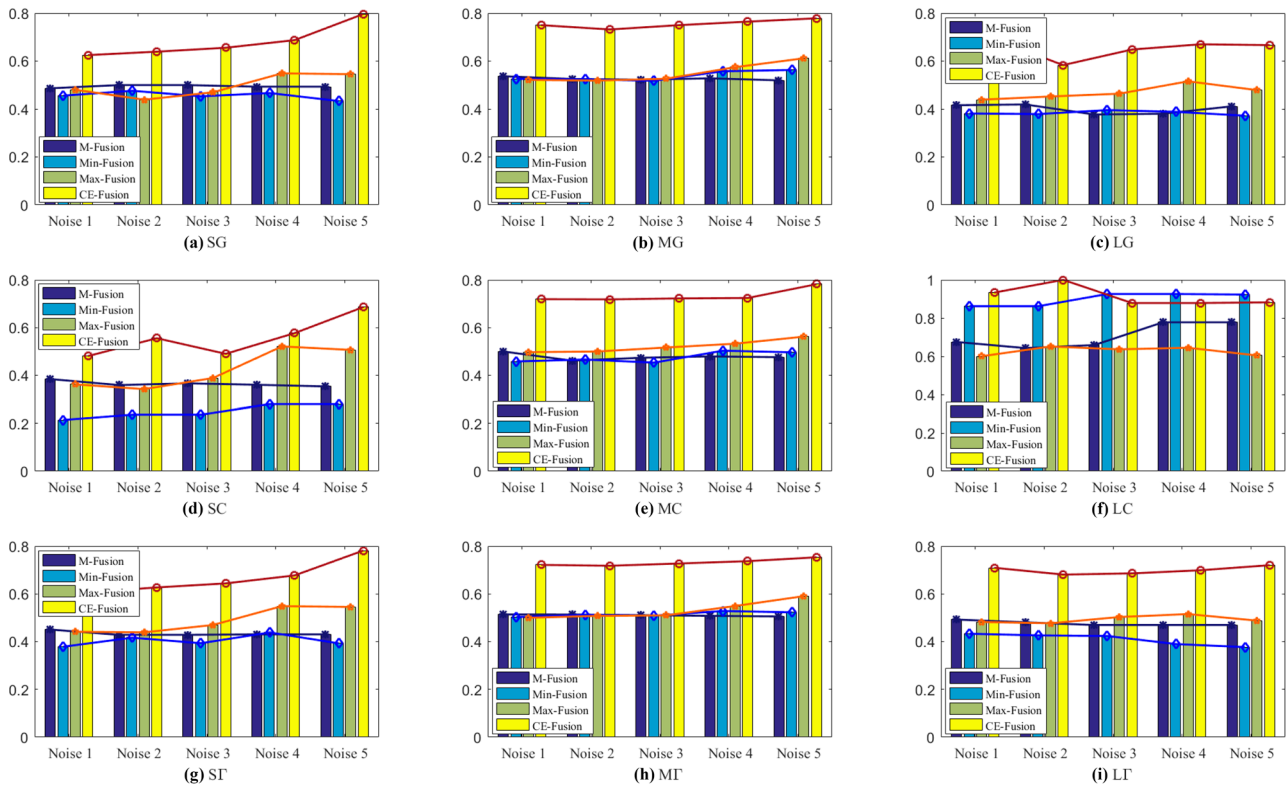


Fig. 4 The approximation accuracy of User Knowledge Modeling under nine noises

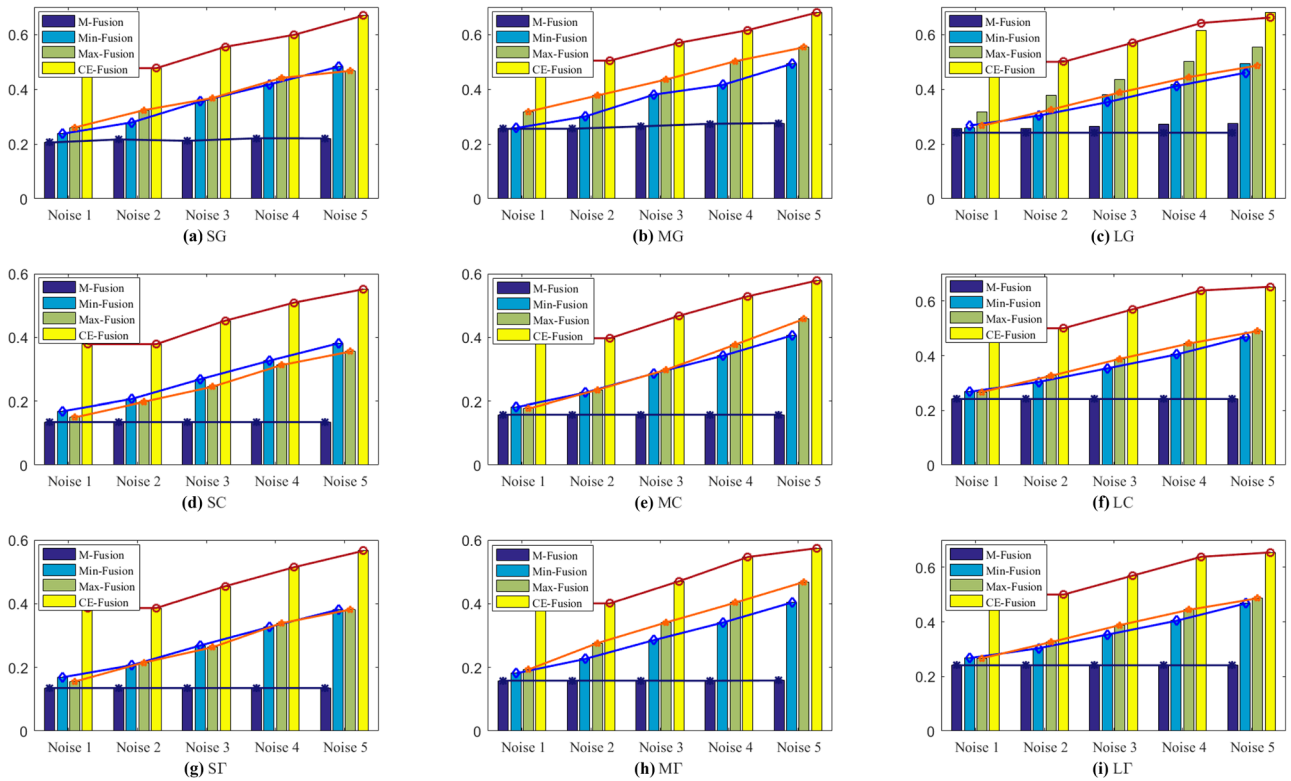


Fig. 5 The approximation accuracy of Balance under nine noises

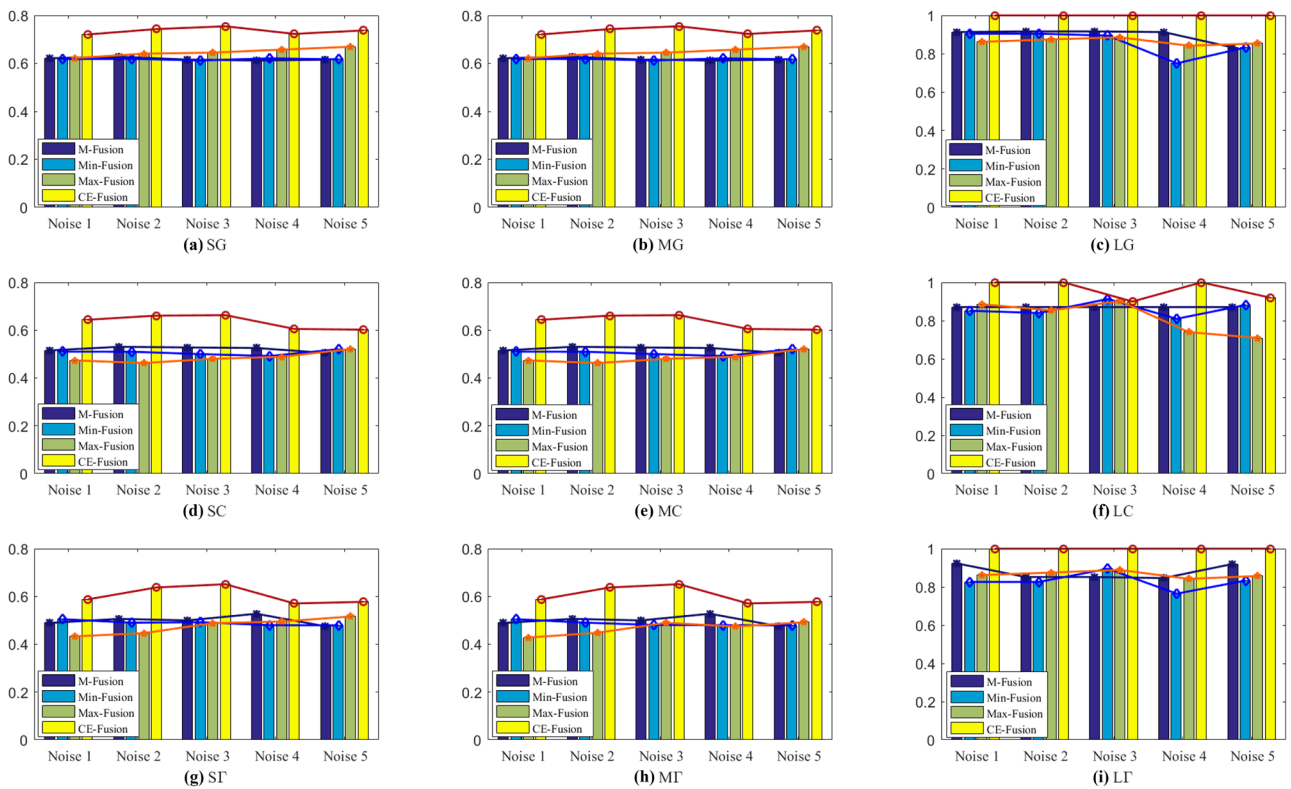


Fig. 6 The approximation accuracy of Pima under nine noises

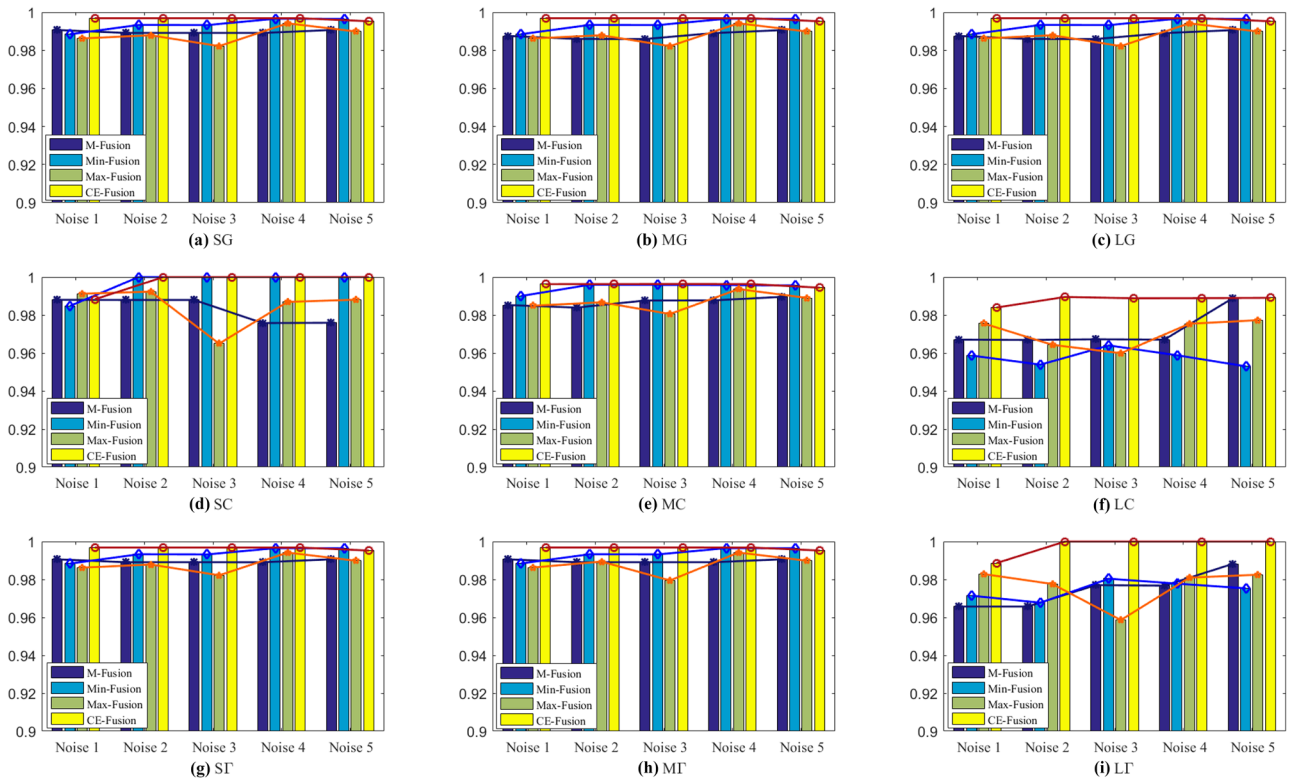


Fig. 7 The approximation accuracy of Vehicle under nine noises

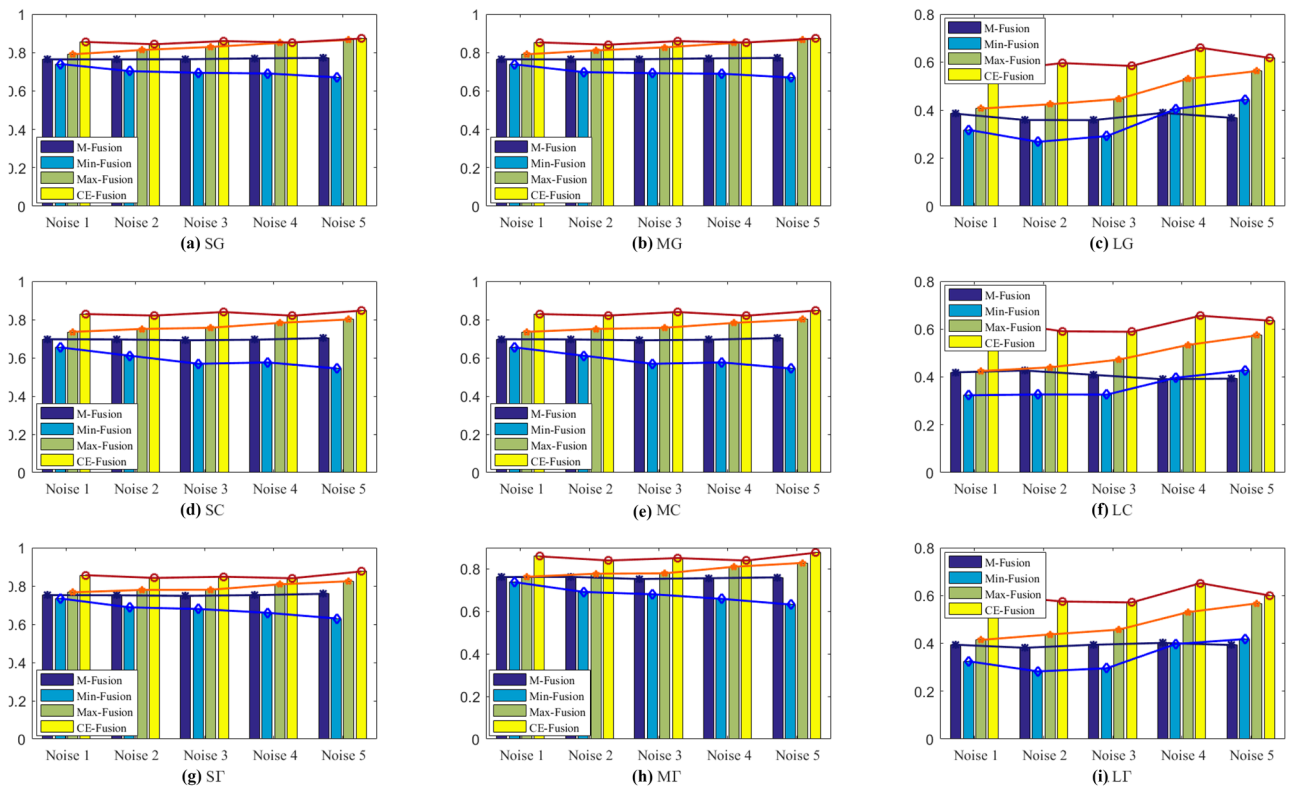


Fig. 8 The approximation accuracy of Winequality_red nine noises

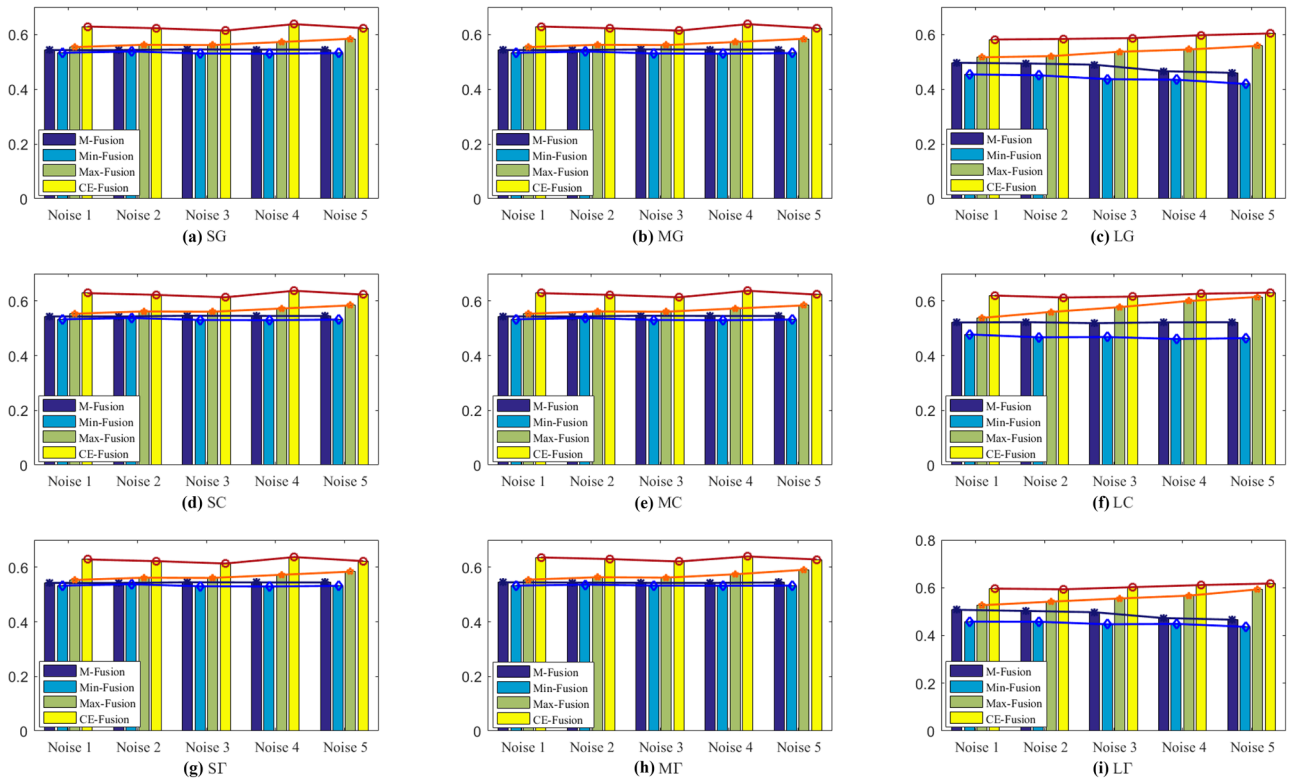


Fig. 9 The approximation accuracy of Wifi_localization_ok Objects under nine noises

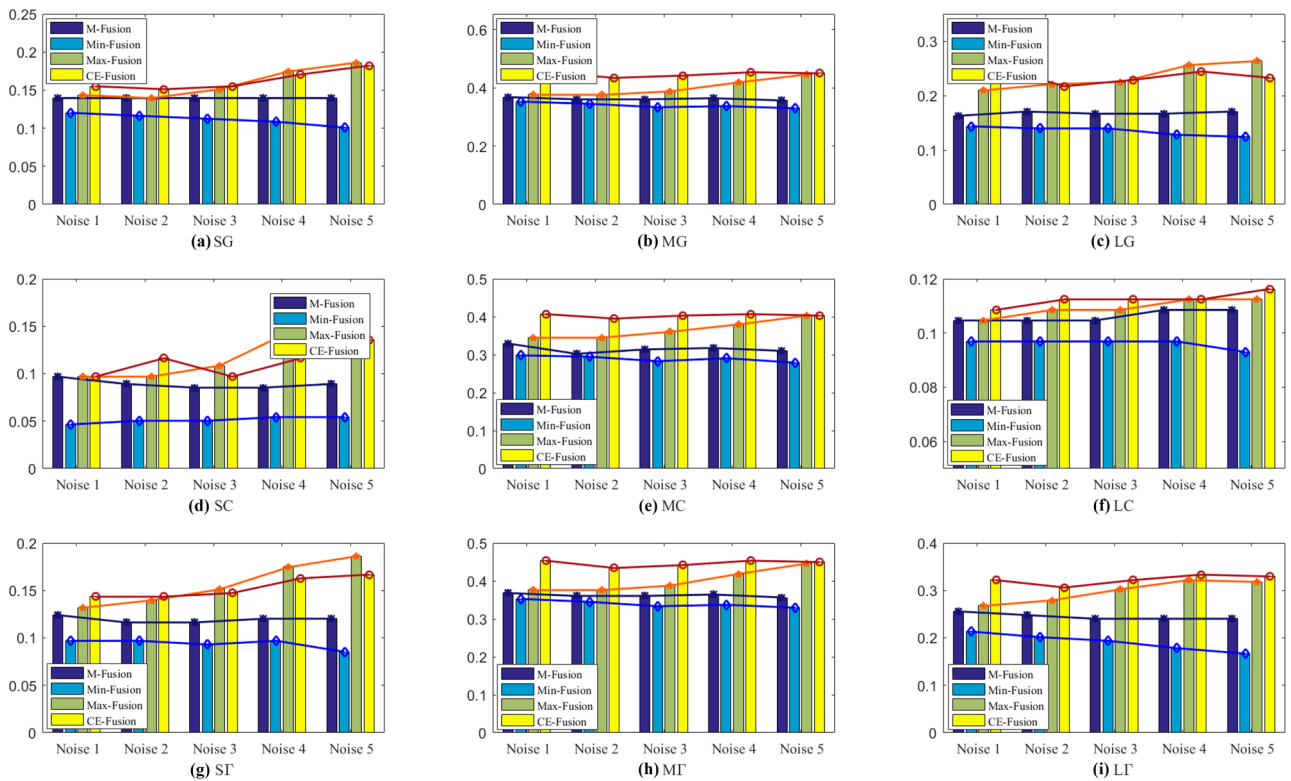


Fig. 10 The approximation quality of User Knowledge Modeling under nine noises

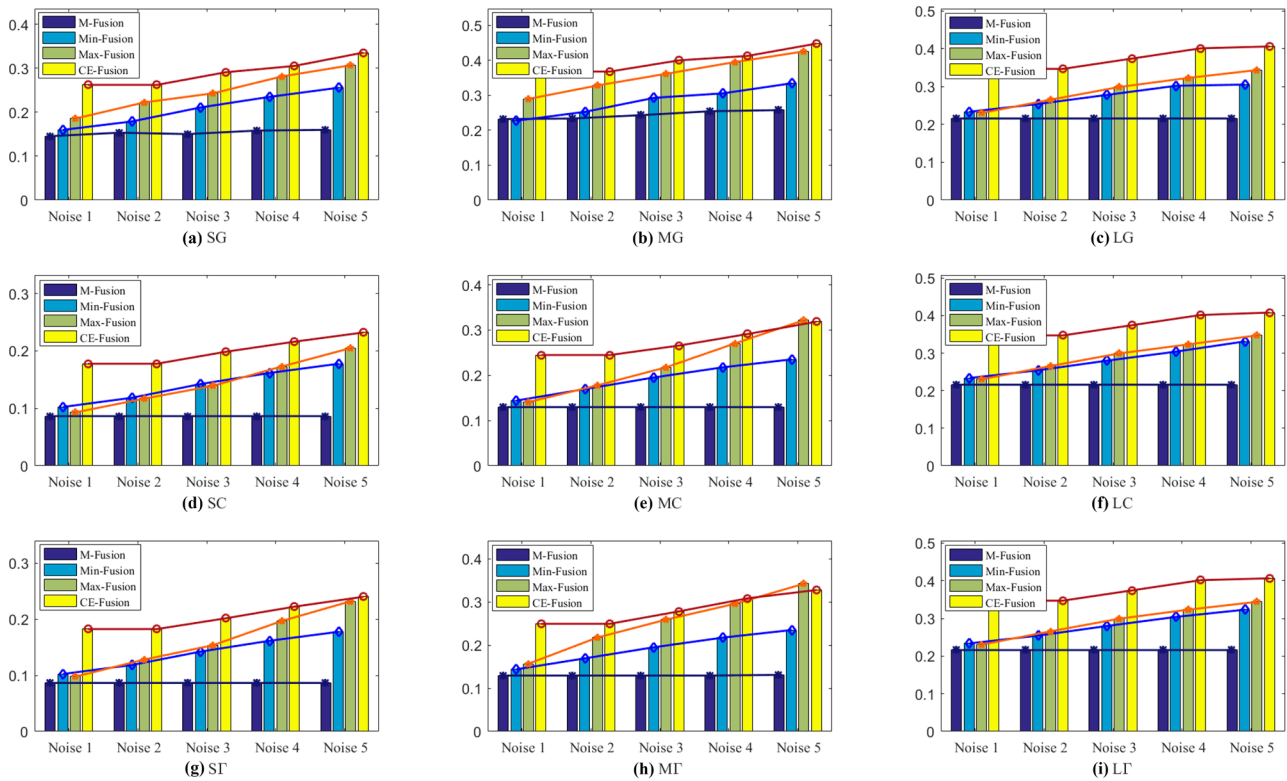


Fig. 11 The approximation quality of Balance under nine noises

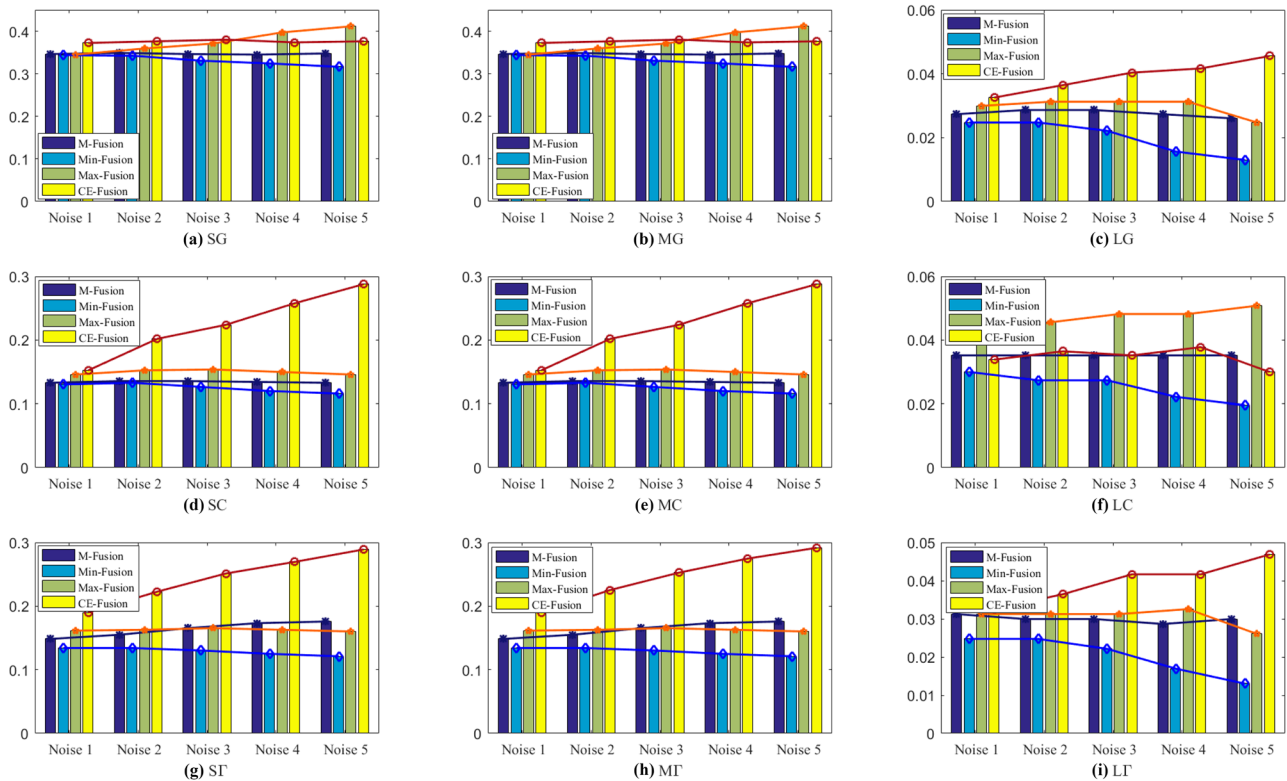


Fig. 12 The approximation quality of Pima under nine noises

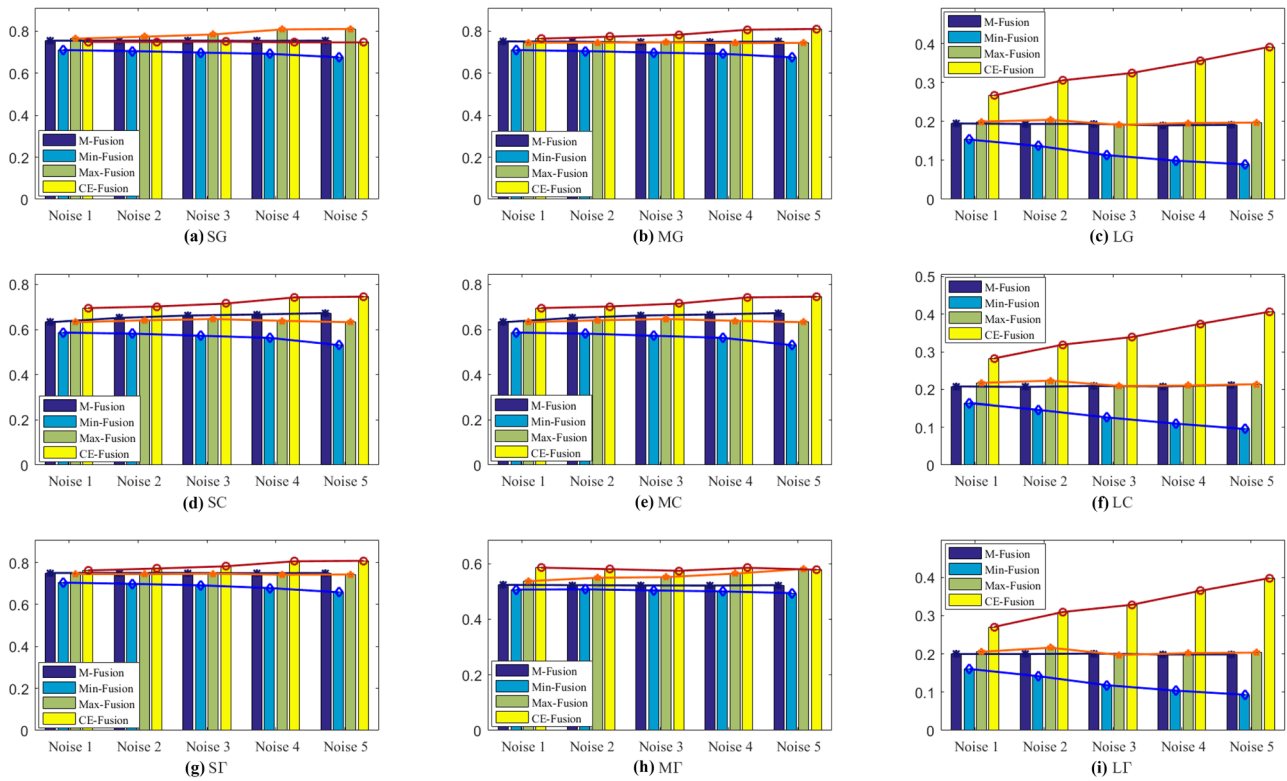


Fig. 13 The approximation quality of Vehicle under nine noises

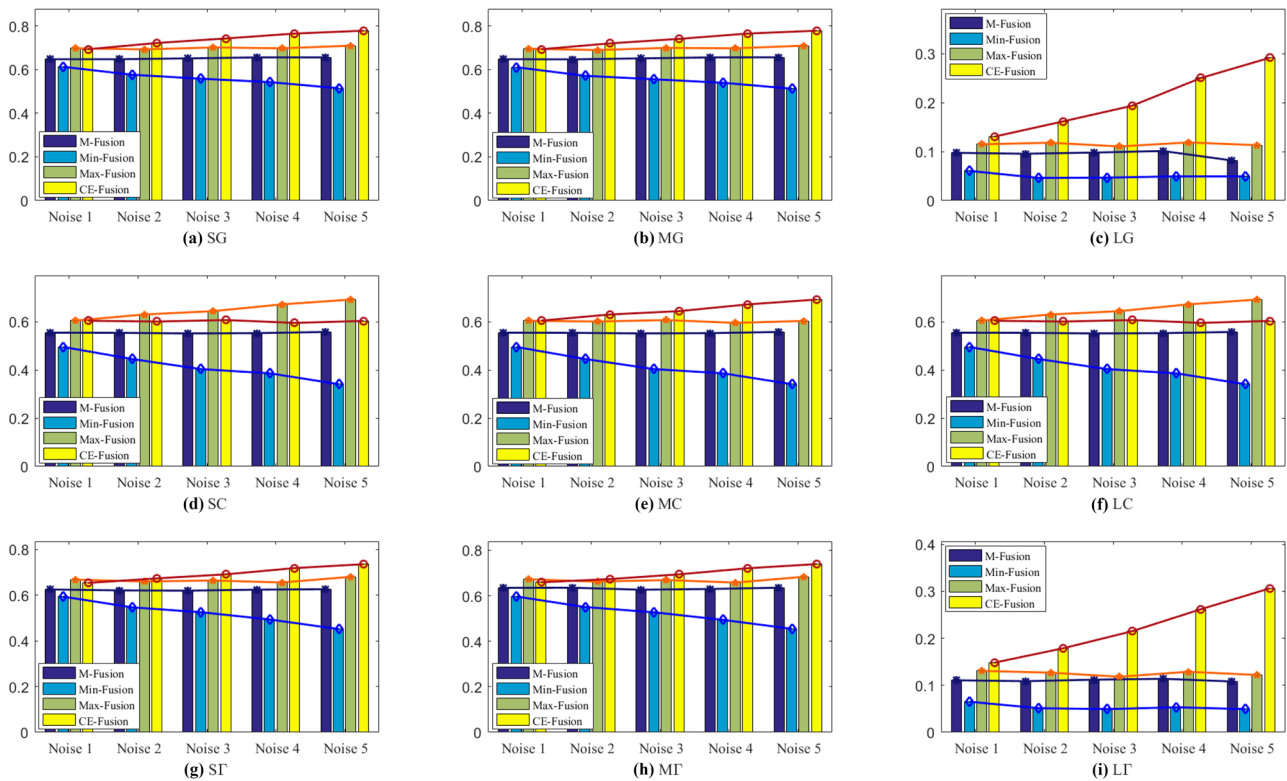


Fig. 14 The approximation quality of Winequality_red nine noises

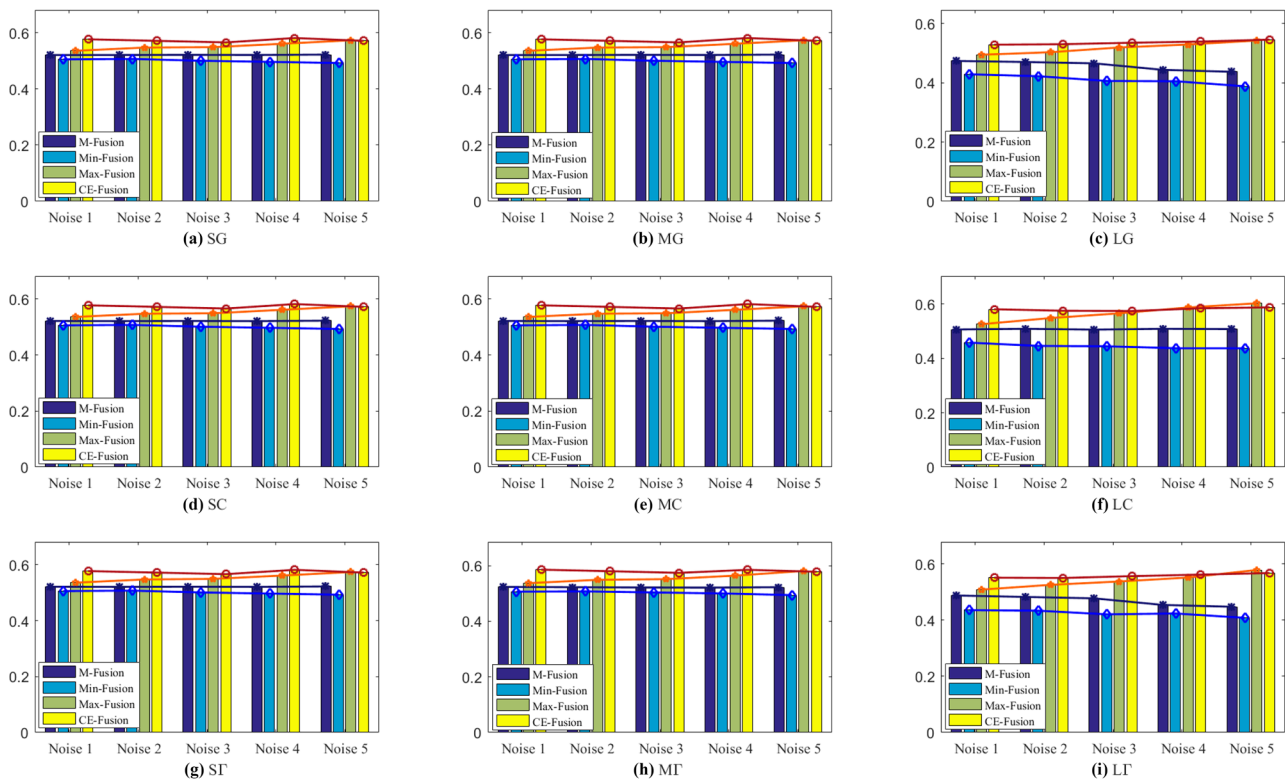


Fig. 15 The approximation quality of Wifi_localization_ok Objects under nine noises

Conclusions

To sum up, concept learning is a challenging, engaging, and promising research direction. Moreover, concept learning deserves to be studied based on granular computing from the perspective of cognitive computing, which may be beneficial to describing and understanding human cognitive processes in a conceptual knowledge way. In order to improve the efficiency and flexibility of concept learning, this paper mainly focuses on concept learning via granular computing from multi-source fuzzy decision tables. It is well known that different fusion methods will have different results when facing some multi-source fuzzy information. The article analyzes the cognitive mechanism of forming concepts from philosophy and cognitive psychology in multi-source fuzzy decision tables. We have considered a novel fusion method based on conditional entropy to fuse multi-source decision tables to describe the cognitive process. Then, granular computing has been combined with the cognitive concept to improve the efficiency of concept learning. The obtained results in this paper may be beneficial to simulating brain intelligence behaviors, including perception, attention, and learning.

Nevertheless, learning cognitive concepts from multi-source information is a challenging task. Although our

work has put forward theoretical frameworks and methods to solve this problem, it is not enough in the application, for example, how to describe approximation of the intent of a cognitive concept in multi-source information tables. It includes the logical and semantic explanation of approximate cognitive concepts, theories, and methods of concept learning from an incomplete or multi-source information table and how to apply these theories in the real world. Some in-depth studies like these issues will be investigated in our future work.

Funding This study was funded by the National Natural Science Foundation of China (Nos. 61976245, 61772002) and Chongqing Postgraduate Research and Innovation Project (No. CYB22153).

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Conflict of Interest The authors declare no competing interests.

References

1. Lake BM, Salakhutdinov R, Tenenbaum JB. Human-level concept learning through probabilistic program induction. *Science*. 2015;350:1332–8.
2. Yao YY. Interpreting concept learning in cognitive informatics and granular computing. *IEEE Trans Syst Man Cybern*. 2009;39:855–66.
3. Li JH, Mei CL, Xu WH, et al. Concept learning via granular computing: a cognitive viewpoint. *Inf Sci*. 2015;298:447–67.
4. Yuan KH, Xu WH, Li WT, et al. An incremental learning mechanism for object classification based on progressive fuzzy three-way concept. *Inf Sci*. 2021;584:127–47.
5. Xu WH, Pang JZ, Luo SQ. A novel cognitive system model and approach to transformation of information granules. *Int J Approx Reason*. 2014;55:853–66.
6. Ganter B. Formal concept analysis-mathematical foundations. Berlin: Springer-Verlag; 1999.
7. Qian T, Wei L, Qi JJ. Constructing three-way concept lattices based on apposition and subposition of formal contexts. *Knowl Based Syst*. 2017;116:39–48.
8. Wille R. Restructuring lattice theory: an approach based on hierarchies of concepts. *Lecture Notes in Artif Intell*. 1982;83:445–70.
9. Wang Y. On concept algebra: a denotational mathematical structure for knowledge and software modelling. *Int J Cogn Inform Natural Intell*. 2008;2:1–19.
10. Wan Q, Li JH, Wei L. Optimal granule combination selection based on multi-granularity triadic concept analysis. *Cogn Comput*. 2021. <https://doi.org/10.1007/s12559-021-09934-6>.
11. Demirkan H, Earley S, Harmon RR. Cognitive computing. *It Professional*. 2017;19:16–20.
12. Shi Y, Mi YL, Li JH, et al. Concurrent concept-cognitive learning model for classification. *Inf Sci*. 2019;496:65–81.
13. Li JH, Huang CC, Qi JJ, et al. Three-way cognitive concept learning via multi-granularity. *Inf Sci*. 2017;378:244–63.
14. Liu PD, Wu Q, Mu XM, et al. Detecting the intellectual structure of library and information science based on formal concept analysis. *Scientometrics*. 2015;104:737–62.
15. Xu WH, Li WT. Granular computing approach to two-way learning based on formal concept analysis in fuzzy dataset. *IEEE Trans Cybern*. 2016;46:366–79.
16. Zhang WX, Xu WH. Cognitive model based on granular computing. *Chinese J Eng Math*. 2007;24:957–71.
17. O’Leary DE. Artificial intelligence and big data. *IEEE Intell Syst*. 2013;28:96–9.
18. Xing EP, Ho Q, Dai W, et al. Petuum: a new platform for distributed machine learning on big data. *IEEE Trans on Big Data*. 2015;1:49–67.
19. Guo DD, Jiang CM, Sheng RX, et al. A novel outcome evaluation model of three-way decision: a change viewpoint. *Inf Sci*. 2022;607:1089–110.
20. Guo DD, Jiang CM, Wu P. Three-way decision based on confidence level change in rough set. *Int J Approx Reason*. 2022;143:57–77.
21. Wu WZ, Leung Y, Mi JS. Granular computing and knowledge reduction in formal contexts. *IEEE Trans Knowl Data Eng*. 2009;21:461–1474.
22. Yao JT, Vasilakos AV, Pedrycz W. Granular computing: perspectives and challenges. *IEEE Trans Cybern*. 2013;43:1977–89.
23. Xu WH, Chen YQ. Multi-attention concept-cognitive learning model: a perspective from conceptual clustering. *Knowl Based Syst*. 2022;252:109472.
24. Yan EL, Yu CG, Lu LM, et al. Incremental concept cognitive learning based on three-way partial order structure. *Knowl Based Syst*. 2021;220:68–98.
25. Zhang T, Rong M, Shan HR, et al. Stability analysis of incremental concept tree for concept cognitive learning. *Int J Mach Learn Cybern*. 2021. <https://doi.org/10.1007/s13042-021-01332-6>.
26. Dai JH, Wang WT, Xu Q. An uncertainty measure for incomplete decision tables and its applications. *IEEE Trans Cybern*. 2013;43:1277–89.
27. Pawlak Z. Rough sets. *International Journal of Computer and Information Sciences*. 1982;11:341–56.
28. Zadeh LA. Fuzzy sets and information granularity. *Advances in Fuzzy Sets Theory and Applications*. 1996: 3-18.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.