Applied Intelligence
https://doi.org/10.1007/s10489-022-03726-7

Uncover the reasons for performance differences between measurement functions (Provably)

Chao Wang[1] · **Jianchuan Feng**[1] · **Linfang Liu**[1] · **Sihang Jiang**[1] · **Wei Wang**[1]

Accepted: 6 May 2022

Abstract
Recently, an exciting experimental conclusion in Li et al. (Knowl Inf Syst 62(2):611–637, 2020) about measures of uncertainty for knowledge bases has attracted great research interest for many scholars. However, these efforts lack solid theoretical interpretations for the experimental conclusion. The main limitation of their research is that the final experimental conclusions are only derived from experiments on three datasets, which makes it still unknown whether the conclusion is universal. In our work, we first review the mathematical theories, definitions, and tools for measuring the uncertainty of knowledge bases. Then, we provide a series of rigorous theoretical proofs to reveal the reasons for the superiority of using the knowledge amount of knowledge structure to measure the uncertainty of the knowledge bases. Combining with experiment results, we verify that knowledge amount has much better performance for measuring uncertainty of knowledge bases. Hence, we prove an empirical conclusion established through experiments from a mathematical point of view. In addition, we find that for some knowledge bases that cannot be classified by entity attributes, such as ProBase (a probabilistic taxonomy), our conclusion is still applicable. Therefore, our conclusions have a certain degree of universality and interpretability and provide a theoretical basis for measuring the uncertainty of many different types of knowledge bases, and the findings of this study have a number of important implications for future practice.

Keywords Concept structure · Knowledge structure · Knowledge base · ProBase · Rough set theory · Uncertainty

1 Introduction

Although knowledge constitutes our area of interest and the cognitive world, it does not have a unified and clear definition [2], which means that knowledge has uncertainty. Uncertainty, including randomness, vagueness, inconsistency, fuzziness, and incompleteness, exists in almost every system and model [3–5], the KBs are no exception. Uncertainty is really a key ingredient in the decision and a fundamental part in modelling [6], therefore, uncertainty is an important research topic in many real-world applications, such as decision making [7], recommendation system [8], Dempster-Shafer evidence theory [9], graph data [10], social networks [11, 12], multi-objective optimization problems [13] and risk analysis during the outbreak of COVID-19 [14–17].

In machine learning tasks, data is an indispensable resource for any machine learning model. However, any machine learning model always has uncertainty when it performs the task of predicting unobserved data. For the KBs, when using the existing knowledge in the KBs to perform inference and decision-making tasks, the uncertainty of the KBs will affect the prediction results of some downstream tasks of natural language understanding. An important reason is the existence of *soft concepts*, which have imprecision. For instance, in the

✉ Chao Wang
cwang17@fudan.edu.cn

Jianchuan Feng
jcfeng20@fudan.edu.cn

Linfang Liu
liulf19@fudan.edu.cn

Sihang Jiang
tedsihangjiang@gmail.com

Wei Wang
wangwei1@fudan.edu.cn

[1] Shanghai Key Laboratory of Data Science, School of Computer Science, Fudan University, Shanghai, China

Published online: 20 June 2022

phrase "`large area`", the definition of `large` lacks strict quantitative standards.

Therefore, how to measure the uncertainty of a system plays a vital role in machine learning, data analysis, artificial intelligence applications, and cognitive science [6]. The current mainstream method is to use **r**ough **s**et **t**heory (RST) [18] to measure the uncertainty of KBs [1, 19]. RST, as a powerful tool that effectively measures the uncertainty of KBs, has attracted more and more attention from artificial intelligence practitioners, such as decision-making [20, 21], computer-aided diagnosis [22], attribute reduction [23], decision analysis [24, 25], and predicting the COVID-19 cases [26]. There are significant advantages in measuring the uncertainty of KBs based on RST. For instance, the RST uses the existing knowledge in the KBs to approximately characterize the unknown knowledge (i.e., target concept) that needs to be explored. The *upper* and *lower approximation* concepts in RST can well describe the uncertainty of KBs [18], and it can be combined with information theory to establish a connection between knowledge uncertainty and information entropy [27]. In addition, the RST is closely related to fuzzy mathematics, which uses the method of describing the fuzziness to measure the uncertainty of knowledge [7, 28].

1.1 Motivation

Based on RST, a series of measurement methods used to measure the uncertainty of the KBs are proposed. For instance, measurement based on the combination of information entropy and rough sets [29]; Using rough entropy theory to measure the uncertainty of KBs [30]; Measurement based on the combination of knowledge granulation and rough sets [31, 32]. Especially in recent work, many scholars focus on the method based on *knowledge structure* [33] to measure the uncertainty of knowledge bases [1, 19]. And obtain many exciting conclusions through a lot of experiments. Although the use of RST to measure the uncertainty of the KBs has achieved a series of great progress, we find that there are still many issues that have not been completely solved.

1. *Conclusions are often based on the verification of a limited number of data sets, lacking a solid and comprehensive theoretical guarantee.* For example, recently, an exciting experimental conclusion in [1] about measures of uncertainty for the KBs has attracted great research interest for scholars. In [1], the authors select three data sets and conduct numerical experiments on these three data sets to verify the superiority of using the knowledge amount to measure the uncertainty of the KBs.[1] However, these successful conclusions lack perfect mathematical expression and interpretability.
2. *The classification of the instances of the knowledge base heavily depends on its attributes.* Using RST to measure the uncertainty of a KB, an important prerequisite is that this KB can be divided by equivalence relations. Unfortunately, subject to certain real task scenarios, some KBs are difficult to meet this condition. For some special datasets, such as ProBase [34], it does not contain a large number of attributes of instances. Therefore, in ProBase, it is difficult to perform the above classification operations on instances based on their attributes. This requires us to transfer the opinions in RST to ProBase for analogy research.

To address the first issue, we employ RST as the theoretical basis to analyze the differences between different methods used to measure the uncertainty in the KBs. Specifically, (1) In terms of theoretical analysis, we compare and analyze in detail the mathematical principles of using *knowledge granulation* of knowledge structure, *knowledge entropy* of knowledge structure, *rough entropy* of knowledge structure and *knowledge amount* of knowledge structure (four measurement functions in total) to measure the uncertainty of the KBs. We find that the above four measurement functions can be unified into an elementary function $\lambda(\cdot)$ (i.e., (12)). The four measurement functions correspond to the four different inputs of function $\lambda(\cdot)$. Based on it, we theoretically prove that the conclusion in [1] is universal and interpretable, and further improved the theory of measures of uncertainty for the KBs. (2) In terms of experimental evaluation, we conduct experiments on 18 public datasets in different fields. The experimental results fully verified our theoretical analysis conclusions.

To address the second issue, we transfer the method of using RST to measure the uncertainty of the KBs to the study of the uncertainty of ProBase. (1) In terms of theoretical analysis, we explore the theoretical feasibility of using RST to measure the uncertainty of ProBase. From the view of RST, equivalence relations determine the partitions on the set $\mathcal{W}$, and get equivalence classes under different equivalence relations thereby. Inspired by this, we regard an equivalence relation in the KBs as a *hypernym* (or *concept*) in ProBase, then we use hypernyms (or concepts) to divide instances, to obtain the equivalence class thereby.

[1] This can be simply understood as knowledge amount has much better performance for measuring uncertainty of knowledge bases, and "performance" can be quantified by objective statistical indicators such as *coefficient of variation*.

To this end, we provide a strategy for inducing datasets from ProBase, and the instances in the induced datasets can be divided by their concepts. (2) In terms of experimental evaluation, in order to verify the above ideas, we induce three datasets based on the strategy in ProBase, and perform experimental verification on three data sets. The experimental results fully verified our theoretical analysis conclusions.

1.2 Contribution

In brief, the contributions in this paper are summarized as follows:

1. We rigorously explain why **k**nowledge **am**ount (KAM) has much better performance for measuring the uncertainty of KBs. We prove an empirical conclusion established through experiments from a mathematical point of view.
2. We prove that measurement methods based on knowledge granulation, knowledge entropy, rough entropy, and knowledge amount can be integrated into a unified measurement function in measuring the uncertainty of KBs. We provide a formal representation of the unified measurement framework and exhaustive comparative analysis.
3. We propose an efficient strategy that induces a new dataset from ProBase. The instances in the induced dataset can be rigorously partitioned based on their concepts. Therefore, we expand the usage scenarios of the measurement function so that the measurement function is still valid for datasets that do not have enough attributes.

1.3 Paper organization

In Section 2, we briefly review the previous studies related to the work of this paper. In Section 3, we review some definitions related to RST, KBs and summarize some notations used in our work. In Section 4, we summarize the calculation methods and properties of the four measurement functions used to measure the uncertainty of KBs. In Section 5, we review the dispersion analysis of numerical experiments in [1]. In Section 6, we conduct a detailed theoretical analysis of different measurement functions and provide our main conclusions (i.e., Theorems 1,2, 3, and 4). Specifically, we unified the four popular measurement functions into a new measurement function. In Section 7, we first provide the definition of the *concept structure* of ProBase (see Definition 13). And then, we provide an effective strategy to induce KBs from ProBase, and instances in induced KBs can be classified by their concept of them. In Section 8, we verify our theoretical analysis via extensive experiments. Specifically, we conduct experiments on 18 public datasets and on three datasets induced from ProBase based on our proposed strategy. Section 11 summarizes our work.

2 Related work

In recent years, research on KBs has become one of the important topics in industry and academia. Many researchers have made exceptional contributions to this field and achieved a series of important results. Especially in theoretical research on the KBs, a series of important results have been obtained. These important conclusions have far-reaching significance for establishing a computable and measurable framework in the KBs. In particular, the uncertainty measurement of KBs based on knowledge structure has been widely concerned.

Knowledge structure Qian et al. [35] describe the differences between various knowledge structures in the KBs based on the concept of knowledge distance. Li et al. [33] propose the definition of lattice, mapping, soft characterizations, and the group of knowledge structures. In the study of the relationship between different KBs, Li et al. [36] regard the KBs as a special relation information system. By introducing homomorphisms, they prove that the KBs are invariant under homomorphisms. Subsequently, based on the homomorphism relation between KBs, Qin et al. [37] propose the concept of communication between KBs, and they obtain a series of invariant characterizations under homomorphisms. It is worth noting that the above works all involve RST, which also provides a strong theoretical basis for our work. In addition, some scholars are committed to using other means to describe the knowledge structure, such as using fuzzy skill maps [38] and knowledge space theory [39].

Measurement method The uncertainty of the KBs is usually calculated by entropy (e.g., information entropy) [40]. Some scholars have shown an increased interest in the combination of entropy theory and rough theory to measure the uncertainty of the system. Hence, many classic mathematical tools have been proposed. For example, Düntsch and Gediga et al. [29] study measuring uncertainty of rough sets with information entropy; Beaubouef et al. [30] propose a new concept, called rough entropy; Liang et al. [27] establish the relationships between rough and information entropy. In the study of knowledge granulation, Wierman [31] focuses on using knowledge granulation to measure the uncertainty of rough sets; Yao [41] employs the concept of granularity measure when studying the probabilistic approaches to rough sets; Shah et al. [32] propose

many measures using soft rough covering sets theory and applied this theory to the task of multi-criteria decision making. Qin et al. [42] use rough set theory to analyze knowledge structures in a tolerance knowledge base. Kobren et al. [43] provide a new framework that can use user feedback to realize the construction and maintenance of the knowledge base in the case of identity uncertainty. Guo and Xu [7] provide a novel entropy-independent measurement function to capture the features of intuitionistic fuzzy sets.

3 Preliminaries

In this section, the key mathematical notations and their descriptions are listed in Table 1, and some basic definitions are reviewed.

Definition 1 ([1] Binary relation $\mathbf{R}$ on $\mathcal{W}$) Let $w_i\mathbf{R}w_j$ denote the binary relation between w_i and w_j on $\mathcal{W}$, where w_i is the *predecessor* of w_j, and w_j is the *successor* of w_i. If $(w_i, w_j) \in \mathbf{R} \subseteq \mathcal{W} \times \mathcal{W}$, then we have $w_i\mathbf{R}w_j$.

For any (w_i, w_j), the binary relation $\mathbf{R}$ can be represented by a 0-1 square matrix as follows,

$$\text{Matrix}(\mathbf{R}) = \begin{bmatrix} \mathbf{R}(w_1, w_1) & \cdots & \mathbf{R}(w_1, w_m) \\ \vdots & \ddots & \vdots \\ \mathbf{R}(w_m, w_1) & \cdots & \mathbf{R}(w_m, w_m) \end{bmatrix}_{m\times m}$$

where $\mathbf{R}(w_i, w_j) = 1$, if $(w_i, w_j) \in \mathbf{R}$, otherwise, $\mathbf{R}(w_i, w_j) = 0$.

Definition 2 ([1, 44] Equivalence relation on $\mathcal{W}$) If $\mathbf{R}$ satisfies the following three properties, then we call $\mathbf{R}$ to be an equivalence relation on $\mathcal{W}$. Specifically,

1. *reflexive* means that $w\mathbf{R}w$ always holds for any $w \in \mathcal{W}$,
2. *symmetric* means that $w\mathbf{R}v$ implies $v\mathbf{R}w$ for any $w, v \in \mathcal{W}$,
3. *transitive* refers to $w\mathbf{R}v$ and $v\mathbf{R}u$ imply $w\mathbf{R}u$ for any $w, u, v \in \mathcal{W}$.

Since $\mathcal{W}$ can be partitioned by an equivalence relation $\mathbf{R}_i$, and the following definition of the equivalence class is obtained.

Definition 3 ([44] Equivalence class on $\mathcal{W}$) Let $\mathbf{R}_i$ be an equivalence relation on $\mathcal{W}$, we call that

$$[w]_{\mathbf{R}_i} = \{v \in \mathcal{W} \mid w\mathbf{R}_i v\}, \tag{1}$$

is the equivalence class including w, and

$$\mathcal{W}/\mathbf{R}_i = \left\{[w]_{\mathbf{R}_i} \mid w \in \mathcal{W}\right\} \tag{2}$$

is the family of all $[w]_{\mathbf{R}_i}$.

Definition 4 ([18] Knowledge base) $[\mathcal{W}, \mathcal{R}]$ is called a KB if and only if $\mathcal{R} \in 2^{\mathcal{R}[\mathcal{W}]}$.

Definition 5 ([44] Equivalence relationship between KBs) Given two KBs $[\mathcal{W}, \mathcal{Q}]$ and $[\mathcal{W}, \mathcal{O}]$, if $[\mathcal{W}, \mathcal{Q}]$ and $[\mathcal{W}, \mathcal{O}]$ are equivalent (i.e., $[\mathcal{W}, \mathcal{Q}] \triangleq [\mathcal{W}, \mathcal{O}]$) then we have

$$[\mathcal{W}, \mathcal{Q}] \triangleq [\mathcal{W}, \mathcal{O}] \Longleftrightarrow \mathcal{W}/\mathcal{Q} \triangleq \mathcal{W}/\mathcal{O}.$$

Table 1 Key Notations and Descriptions

Notation	Description
$\emptyset$	the empty set
$\mathbb{R}$	the set of real numbers
$\mathbb{Z}^+$	the set of positive integers
$\mathcal{W}$	a non-empty finite set, named *universe* $\mathcal{W}$
$2^{\mathcal{W}}$	the family of all subsets of $\mathcal{W}$
$w_i\mathbf{R}w_j$	the binary relation between w_i and w_j on $\mathcal{W}$
$\mathcal{R} = \{\mathbf{R}_i\}_{n_1}$	the set of all binary relations $\mathbf{R}_i$ on universe $\mathcal{W}$
$\mathcal{O} = \{\mathbf{O}_i\}_{n_2}$	the set of all binary relations $\mathbf{O}_i$ on universe $\mathcal{W}$
$\mathcal{P} = \{\mathbf{P}_i\}_{n_3}$	the set of all binary relations $\mathbf{P}_i$ on universe $\mathcal{W}$
$\mathcal{Q} = \{\mathbf{Q}_i\}_{n_4}$	the set of all binary relations $\mathbf{Q}_i$ on universe $\mathcal{W}$
$\mathcal{R}[\mathcal{W}]$	the family of all equivalence relations on $\mathcal{W}$
$\lvert W\rvert$	the cardinality of W, e.g., $\lvert\{a, b, c\}\rvert = 3$
$M \triangleq N$	M and N are equivalent, where M and N be two functions or sets
$\mathcal{W} = \{w_i\}_k$	the simplified form of $\mathcal{W} = \{w_1, w_2, ..., w_k\}$
$[\mathcal{W}, \mathcal{R}]$	the knowledge base
$[\mathcal{T}, \mathcal{H}]$	the knowledge base induced by ProBase
$M(\mathcal{W})$	the measure set on $\mathcal{W}$

Definition 6 ([1] Knowledge structure of $[\mathcal{W}, \mathcal{R}]$) If the finite set $\mathcal{W} = \{w_i\}_k$ can be divided by relations $\mathcal{R} = \{\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_i\}$, then we call the vector

$$\mathrm{CSV}(\mathcal{R}) = \left\langle [w_1]_{\mathcal{R}}, [w_2]_{\mathcal{R}}, \ldots, [w_k]_{\mathcal{R}} \right\rangle \tag{3}$$

the knowledge structure of $[\mathcal{W}, \mathcal{R}]$.

Definition 7 (Indiscernibility relation over $\mathcal{P}$) If $\emptyset \neq \mathcal{P} \subseteq \mathcal{R}$, then we call $\bigcap \mathcal{P}$ is the indiscernibility relation over $\mathcal{P}$, which is denoted by $ind(\mathcal{P})$.

In other words, let F be the finite set, and f_a and f_b are two entities in F. f_a and f_b satisfy *indiscernibility relation* over $\mathcal{P}$ if and only if f_a and f_b have the same value on all elements in $\mathcal{P}$. For example, a *red* `Porsche` and a *red* `Tesla` satisfy the indiscernibility relation on the attribute *color*.

Example 1 Given a collection $\mathcal{W} = \{w_1, w_2, \cdots, w_8\}$ that contains 8 `candies`. Suppose these `candies` have different *colors* (e.g., *red*, *blue*, *yellow*), *shapes* (e.g., *square*, *round*, *triangular*), *flavors* (e.g., *lemony*, *sweet*). Therefore, these `candies` can be divided according to *color*, *shape* and *taste*. Statistical information about $\mathcal{W}$ is summarized in Table 2.

As shown in Table 2, we can define three equivalence relations, namely, R_1 (i.e., *color*), R_2 (i.e., *shape*), and R_3 (i.e., *taste*). Further, through these three equivalence relations, the following three equivalence classes are obtained, i.e.,

$$\begin{aligned}
\mathcal{W}/R_1 &= \{\{w_1, w_3, w_7\}, \{w_2, w_4\}, \{w_5, w_6, w_8\}\},\\
\mathcal{W}/R_2 &= \{\{w_1, w_5\}, \{w_2, w_6\}, \{w_3, w_4, w_7, w_8\}\},\\
\mathcal{W}/R_3 &= \{\{w_2, w_3, w_7\}, \{w_1, w_3, w_4, w_5, w_6\}\}.
\end{aligned}$$

Apparently, according to Definition 4, $[\mathcal{W}, \{R_1, R_2, R_3\}]$ is the KB. And according to Definition 7, w_1 and w_3 satisfy the indiscernibility relation on the color *red*, w_1 and w_4 satisfy the indiscernibility relation on the shape *square*.

Table 2 Candies are divided according to *color*, *shape* and *taste*

Attribute	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
Red	✓		✓				✓	
Blue		✓		✓				
Yellow					✓	✓		✓
Square	✓			✓				
Round		✓				✓		
Triangular			✓	✓			✓	✓
Lemony		✓					✓	✓
Sweet	✓		✓	✓	✓	✓		

4 Four uncertainty measurement functions for KBs

In this section, we introduce the categories, the core idea, and the formalization of four measurement functions. It is worth noting that, for a finite set $\mathcal{W}$, we can divide $\mathcal{W}$ based on its equivalence relations $\mathcal{R}$ (based on rough set theory guidance) to obtain the knowledge base $[\mathcal{W}, \mathcal{R}]$. Then, according to Definition 6, we obtain the knowledge structure (i.e., $\mathrm{CSV}(\mathcal{R})$) of $[\mathcal{W}, \mathcal{R}]$. Moreover, based on the $\mathrm{CSV}(\mathcal{R})$, we can unitize the knowledge granulation of $\mathrm{CSV}(\mathcal{R})$, the knowledge entropy of $\mathrm{CSV}(\mathcal{R})$, the rough entropy of $\mathrm{CSV}(\mathcal{R})$, and the knowledge amount of $\mathrm{CSV}(\mathcal{R})$ to construct the measure set, respectively. Finally, based on the constructed measure set (the principles of measure set construction and example are provided in Section 8), and the coefficient of variation (denoted as $C_v(\mathcal{W})$ in (11), which is a common objective statistical indicator used to measure the uncertainty of a dataset) of the set is calculated to measure the uncertainty of the KB $[\mathcal{W}, \mathcal{R}]$.

4.1 Categories of four measurement functions

In this paper, we focus on 4 currently popular measurement functions for measuring uncertainty of knowledge bases. Specifically, these methods include:

1. **Granularity-based measures** (i.e., the knowledge granulation of $\mathrm{CSV}(\mathcal{R})$ in Definition 8).
2. **Entropy-based measures** (i.e., the knowledge entropy of $\mathrm{CSV}(\mathcal{R})$ in Definition 9, and the rough entropy of $\mathrm{CSV}(\mathcal{R})$ in Definition 10).
3. **Knowledge amount-based measures** (i.e., the knowledge amount of $\mathrm{CSV}(\mathcal{R})$ in Definition 11).

4.2 The core idea of four measurement functions

1. **The core idea of granularity-based measures:** The granulation of knowledge in the KB is mainly quantified by counting the number of elements in the equivalence relations $\mathbf{R} \in \mathcal{R}$. Specifically, given a KB $[\mathcal{W}, \mathcal{R}]$, if $\mathcal{R} \in 2^{\mathcal{R}[\mathcal{W}]}$, then the granulation of $[\mathcal{W}, \mathcal{R}]$ can be formalized as a mapping function from $2^{\mathcal{R}[\mathcal{W}]}$ to $(0, +\infty]$.
2. **The core idea of entropy-based measures:** In classical thermodynamics, entropy as a measurable physical property reveals the disorder of the system (the higher the value of entropy, the higher disorder of the system). In information theory, entropy (e.g., Shannon entropy) is used to measure the uncertainty of a system. Similarly, a large number of studies applied the concept of entropy to measure the uncertainty of KBs.

3. **The core idea of knowledge amount-based measures:** These measures are the variation of the entropy-based measures described above, which introduces a probability measure (e.g., the probability of W_i in the universe $\mathcal{W}$). These makes it possible to measure the uncertainty and the fuzziness of the KB.

4.3 Formalization of four measurement functions

Definition 8 [[1] Knowledge granulation of CSV($\mathcal{R}$)] For a knowledge base $[\mathcal{W}, \mathcal{R}]$, the knowledge granulation of CSV($\mathcal{R}$) is quantified as:

$$\mathrm{KGR}(\mathcal{R}) = \frac{1}{k^2}\sum_{i=1}^{m}|W_i|^2 = \frac{1}{k^2}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}}\right|, \tag{4}$$

where $\mathcal{W}/\mathcal{R} = \{W_i\}_m$, $W_i = \{w_i\}_{n_i}$ (i.e., $|W_i| = n_i$), $\sum_{i=1}^{m}(n_i) = \sum_{i=1}^{m}|W_i| = |\mathcal{W}| = k$. $\mathcal{R}$ is the set of equivalence relations.

Definition 9 ([1] Knowledge entropy of CSV($\mathcal{R}$)) For a knowledge base $[\mathcal{W}, \mathcal{R}]$, the knowledge entropy of CSV($\mathcal{R}$) is quantified as:

$$\mathrm{KEN}(\mathcal{R}) = -\sum_{i=1}^{m}\frac{|W_i|}{k}\log_2\frac{|W_i|}{k} = -\sum_{i=1}^{k}\frac{1}{k}\log_2\frac{\left|[w_i]_{\mathcal{R}}\right|}{k} \tag{5}$$

where $\mathcal{W}/\mathcal{R} = \{W_i\}_m$, $W_i = \{w_i\}_{n_i}$ (i.e., $|W_i| = n_i$), $\sum_{i=1}^{m}(n_i) = \sum_{i=1}^{m}|W_i| = |\mathcal{W}| = k$. $\mathcal{R}$ is the set of equivalence relations.

Definition 10 ([1] Rough entropy of CSV($\mathcal{R}$)) For a knowledge base $[\mathcal{W}, \mathcal{R}]$, the rough entropy of CSV($\mathcal{R}$) is quantified as:

$$\mathrm{REN}(\mathcal{R}) = -\sum_{i=1}^{m}\frac{|W_i|}{k}\log_2\frac{1}{|W_i|} = -\sum_{i=1}^{k}\frac{1}{k}\log_2\frac{1}{\left|[w_i]_{\mathcal{R}}\right|} \tag{6}$$

where $\mathcal{W}/\mathcal{R} = \{W_i\}_m$, $W_i = \{w_i\}_{n_i}$ (i.e., $|W_i| = n_i$), $\sum_{i=1}^{m}(n_i) = \sum_{i=1}^{m}|W_i| = |\mathcal{W}| = k$. $\mathcal{R}$ is the set of equivalence relations.

Definition 11 ([1] Knowledge amount of CSV($\mathcal{R}$)) For a knowledge base $[\mathcal{W}, \mathcal{R}]$, the knowledge amount of CSV($\mathcal{R}$) is quantified as:

$$\begin{aligned}\mathrm{KAM}(\mathcal{R}) &= \sum_{i=1}^{m}\frac{1}{k^2}|W_i|\,|\mathcal{W}-W_i| \\ &= \sum_{i=1}^{k}\frac{1}{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}}\right|}{k}\right),\end{aligned} \tag{7}$$

where $\mathcal{W}/\mathcal{R} = \{W_i\}_m$, $W_i = \{w_i\}_{n_i}$ (i.e., $|W_i| = n_i$), $\sum_{i=1}^{m}(n_i) = \sum_{i=1}^{m}|W_i| = |\mathcal{W}| = k$. $\mathcal{R}$ is the set of equivalence relations.

4.4 The main properties of KGR($\mathcal{R}$), KEN($\mathcal{R}$), REN($\mathcal{R}$), and KAM($\mathcal{R}$)

Lemma 1 ([1] Boundedness) *Suppose that* $[\mathcal{W}, \mathcal{R}]$ *is a KB and* $|\mathcal{W}| = k$, *then*

$$\begin{aligned}\tfrac{1}{k} \le\ & KGR(\mathcal{R}) \le 1,\\ 0 \le\ & REN(\mathcal{R}) \le \log_2 k,\\ 0 \le\ & KAM(\mathcal{R}) \le \tfrac{k-1}{k},\\ 0 \le\ & KEN(\mathcal{R}) \le \log_2 k.\end{aligned} \tag{8}$$

Inequalities in (8) reveal the boundedness of KGR($\mathcal{R}$), KEN($\mathcal{R}$), REN($\mathcal{R}$), and KAM($\mathcal{R}$) on $\mathcal{W}$.

Lemma 2 ([1] Monotonicity) *Let* $[\mathcal{W}, \mathcal{O}]$, $[\mathcal{W}, \mathcal{Q}]$ *be two KBs. If* $CSV(\mathcal{O}) \prec CSV(\mathcal{Q})$*(i.e., IDE* $(CSV(\mathcal{O})/CSV(\mathcal{Q})) = 1$*), then*

$$\begin{aligned}KGR(\mathcal{O}) &< KGR(\mathcal{Q}),\\ REN(\mathcal{O}) &< REN(\mathcal{Q}),\\ KAM(\mathcal{O}) &> KAM(\mathcal{Q}),\\ KEN(\mathcal{O}) &> KEN(\mathcal{Q}).\end{aligned} \tag{9}$$

For rigorous proof of Lemma 1 and 2, the reader is referred to [1].

5 Dispersion analysis

In this section, we first review the conclusion of numerical experiments of [1]. The authors construct 4 measure sets (the principles of measure set construction and example are provided in Section 8) on three datasets[2] (Nursery, Solar Flare, and Tic-Tac-Toe Endgamelaintaio in Table 3). Then, they compare the performance of four measurement functions (i.e., Definitions 8-11) by dispersion analysis. In their numerical experiment, they use the *coefficient of variation* of datasets to compare the performance differences between four different measurement functions. The experimental results are shown in Table 3.

According to Table 3, it is easy to see that this may imply an interesting conclusion, i.e.,

$$\begin{aligned}C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) &> C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))\\ &> C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W})).\end{aligned} \tag{10}$$

Inequality (10) shows that KAM($\mathcal{P}_i/\mathcal{O}_i/\mathcal{Q}_i$) has a much better performance. The conclusion of Inequality (10) and

[2] https://archive.ics.uci.edu/ml/datasets.php

Table 3 may reflect a kind of regularity, which naturally leads to further thinking about the following questions:

1. *Does the conclusion of* (10) *apply to most datasets?*
2. *Does* (10) *reveal general laws?*
3. *What is the mathematical principle of* (10)*?*

This motivates us to conduct deeper insight into different measurement functions. In the next section, we will give answers to these three questions.

6 Theoretical analysis of measurement functions

In this section, we answer the above three questions. We provide a unified framework to prove Inequality (10), and theoretically prove that Inequality (10) has general properties for most KBs. These conclusions provide a rigorous theoretical basis for measuring uncertainty for KBs. Before giving the conclusions, we review the mathematical tools and notations we need to use in our proof. Specifically, for a given finite set $\mathcal{W} = \{w_i\}_n$, we use $\sigma(\mathcal{W})$ and $C_v(\mathcal{W})$ to represent *standard deviation* and *coefficient of variation* of $\mathcal{W}$, respectively, i.e.,

$$\bar{w} = \frac{1}{n}\sum_{i=1}^{n} w_i,\ \sigma(\mathcal{W}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(w_i - \bar{w})^2},\ C_v(\mathcal{W}) = \frac{\sigma(\mathcal{W})}{\bar{w}}. \tag{11}$$

Next, we provide our core theorems, which are Theorems 1,2.3, and 4. These conclusions strictly theoretically prove the experimental conclusion in [1], solving the two questions raised in Section 5 thereby.

Theorem 1 *Suppose that* $[\mathcal{W}, \mathcal{R}_n]$ *be a KB. Let* $\boldsymbol{M}(\mathcal{W})$ *be the measure set on* $\mathcal{W}$*, where* $\mathcal{W} = \{w_i\}_k$*, which can be divided by relation* $\mathcal{R}_n = \{\boldsymbol{R}_j\}_n$*. Then the* $C_v(\boldsymbol{M}_{KGR}(\mathcal{W}))$ *can be equivalently described by the measurement function* $\lambda(x)$*, where*

$$\lambda(\cdot) = \frac{\sqrt{n \cdot \sum_{i=1}^{n}\left(\sum_{i=1}^{k}(\cdot) - \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{k}(\cdot)\right)^2}}{\sum_{j=1}^{n}\sum_{i=1}^{k}(\cdot)},$$
$$x = |[w_i]_{\mathcal{R}_j}| \in \mathbb{Z}^+. \tag{12}$$

Proof Suppose that $[\mathcal{W}, \mathcal{R}_n]$ be a KB, and let $\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})$ be the measure set on the $\mathcal{W}$ based on knowledge granulation, we suppose that,

$$\begin{aligned}\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}) &= \{\mathrm{KGR}(\mathcal{R}_1), \mathrm{KGR}(\mathcal{R}_2), ..., \mathrm{KGR}(\mathcal{R}_n)\}\\ &= \{\mathrm{KGR}(\mathcal{R}_j)\}_n.\end{aligned} \tag{13}$$

According to (11), we obtain the following, i.e.,

$$\begin{aligned}\overline{\mathrm{KGR}(\mathcal{R})} &= \frac{1}{n}\sum_{j=1}^{n}\mathrm{KGR}(\mathcal{R}_j),\\ \sigma(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KGR}(\mathcal{R}_j) - \overline{\mathrm{KGR}(\mathcal{R})}\right)^2},\\ C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) &= \frac{\sigma(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))}{\overline{\mathrm{KGR}(\mathcal{R})}}.\end{aligned} \tag{14}$$

According to (4), for the set $\mathcal{W} = \{w_i\}_k$ (i.e., $|\mathcal{W}| = k$), it follows that,

$$\mathrm{KGR}(\mathcal{R}) = \frac{1}{k^2}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|^2, \tag{15}$$

and

$$\begin{aligned}\overline{\mathrm{KGR}(\mathcal{R})} &= \frac{1}{n}\sum_{j=1}^{n}\mathrm{KGR}(\mathcal{R}_j)\\ &= \frac{1}{nk^2}\sum_{j=1}^{n}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|.\end{aligned} \tag{16}$$

Further, we obtain

$$\begin{aligned}&\sigma(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))\\ &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KGR}(\mathcal{R}_j) - \overline{\mathrm{KGR}(\mathcal{R})}\right)^2}\\ &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{k^2}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right| - \frac{1}{nk^2}\sum_{j=1}^{n}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|\right)^2}\end{aligned} \tag{17}$$

Table 3 C_v-values of measure sets $\mathbf{M}$(KGR), $\mathbf{M}$(REN), $\mathbf{M}$(KEN) and $\mathbf{M}$(KAM)

Date set	$C_v(\mathbf{M}(\mathrm{KGR}))$	$C_v(\mathbf{M}(\mathrm{REN}))$	$C_v(\mathbf{M}(\mathrm{KEN}))$	$C_v(\mathbf{M}(\mathrm{KAM}))$
Nursery	2.0431	0.6978	0.4750	0.1141
Solar Flare	0.9857	0.3219	0.2806	0.0615
Tic-Tac-Toe Endgame	1.7882	0.9015	0.4340	0.1186

, and

$$\begin{aligned} & C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) \\ &= \frac{\sigma(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))}{\overline{\mathrm{KGR}(\mathcal{R})}} \\ &= \frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KGR}(\mathcal{R}_j)-\overline{\mathrm{KGR}(\mathcal{R})}\right)^2}}{\frac{1}{nk^2}\sum_{j=1}^{n}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|} \\ &= \frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{k^2}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|-\frac{1}{nk^2}\sum_{j=1}^{n}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|\right)^2}}{\frac{1}{nk^2}\sum_{j=1}^{n}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|} \\ &= \frac{\sqrt{n\cdot\sum_{i=1}^{n}\left(\sum_{i=1}^{k}(|[w_i]_{\mathcal{R}_j}|)-\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{k}(|[w_i]_{\mathcal{R}_j}|)\right)^2}}{\sum_{j=1}^{n}\sum_{i=1}^{k}(|[w_i]_{\mathcal{R}_j}|)}. \end{aligned} \tag{18}$$

By (18), we establish the mapping relationship between $C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))$ and $|[w_i]_{\mathcal{R}_j}|$, i.e.,

$$\lambda(|[w_i]_{\mathcal{R}_j}|) = C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) \triangleq C_v(\mathrm{KGR}(\{\mathcal{R}_j\})_n), \tag{19}$$

where $\lambda(\cdot)$ satisfies (12). The proof is completed. □

Theorem 2 *Suppose that $[\mathcal{W}, \mathcal{R}_n]$ be a KB. Let $\boldsymbol{M}(\mathcal{W})$ be the measure set on $\mathcal{W}$, where $\mathcal{W} = \{w_i\}_k$, which can be divided by relation $\mathcal{R}_n = \{\boldsymbol{R}_j\}_n$. Then the $C_v(\boldsymbol{M}_{REN}(\mathcal{W}))$ can be equivalently described by the measurement function $\lambda(x)$ (i.e., (12)).*

Proof Suppose that $[\mathcal{W}, \mathcal{R}_n]$ be a KB, and let $\mathbf{M}_{\mathrm{REN}}(\mathcal{W})$ be the measure set on the $\mathcal{W}$ based on rough entropy, we suppose that,

$$\begin{aligned} \mathbf{M}_{\mathrm{REN}}(\mathcal{W}) &= \{\mathrm{REN}(\mathcal{R}_1), \mathrm{REN}(\mathcal{R}_2), ..., \mathrm{REN}(\mathcal{R}_n)\} \\ &= \{\mathrm{REN}(\mathcal{R}_j)\}_n. \end{aligned} \tag{20}$$

According to (11), then we obtain the following, i.e.,

$$\begin{aligned} \overline{\mathrm{REN}(\mathcal{R})} &= \frac{1}{n}\sum_{j=1}^{n}\mathrm{REN}(\mathcal{R}_j), \\ \sigma(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{REM}(\mathcal{R}_j)-\overline{\mathrm{REN}(\mathcal{R})}\right)^2}, \\ C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) &= \frac{\sigma(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}))}{\overline{\mathrm{REN}(\mathcal{R})}} \end{aligned} \tag{21}$$

According to (6), for the set $\mathcal{W} = \{w_i\}_k$ (i.e., $|\mathcal{W}| = k$), it follows that,

$$\mathrm{REN}(\mathcal{R}_j) = -\sum_{i=1}^{k}\frac{1}{k}\log_2\frac{1}{\left|[w_i]_{\mathcal{R}_j}\right|} = \sum_{i=1}^{k}\frac{1}{k}\log_2\left|[w_i]_{\mathcal{R}_j}\right|, \tag{22}$$

and

$$\overline{\mathrm{REN}(\mathcal{R})} = \frac{1}{n}\sum_{j=1}^{n}\mathrm{REN}(\mathcal{R}_j) = \frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\left|[w_i]_{\mathcal{R}_j}\right|. \tag{23}$$

Further, we obtain

$$\begin{aligned} & \sigma(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) \\ &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{REN}(\mathcal{R}_j)-\overline{\mathrm{REN}(\mathcal{R})}\right)^2} \\ &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{k}\frac{1}{k}\log_2\left|[w_i]_{\mathcal{R}_j}\right|-\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\left|[w_i]_{\mathcal{R}_j}\right|\right)^2} \end{aligned} \tag{24}$$

, and

$$\begin{aligned} & C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) \\ &= \frac{\sigma(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}))}{\overline{\mathrm{REN}(\mathcal{R})}} \\ &= \frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{REN}(\mathcal{R}_j)-\overline{\mathrm{REN}(\mathcal{R})}\right)^2}}{\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|} \\ &= \frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{k}\frac{1}{k}\log_2\left|[w_i]_{\mathcal{R}_j}\right|-\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\left|[w_i]_{\mathcal{R}_j}\right|\right)^2}}{\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\left|[w_i]_{\mathcal{R}_j}\right|} \\ &= \frac{\sqrt{n\cdot\sum_{i=1}^{n}\left(\sum_{i=1}^{k}(\log_2|[w_i]_{\mathcal{R}_j}|)-\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{k}(\log_2|[w_i]_{\mathcal{R}_j}|)\right)^2}}{\sum_{j=1}^{n}\sum_{i=1}^{k}(\log_2|[w_i]_{\mathcal{R}_j}|)}. \end{aligned} \tag{25}$$

By (25), we establish the mapping relationship between $C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}))$ and $\log_2|[w_i]_{\mathcal{R}_j}|$, i.e.,

$$\lambda(\log_2|[w_i]_{\mathcal{R}_j}|) = C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) \triangleq C_v(\mathrm{REN}(\{\mathcal{R}_j\})_n), \tag{26}$$

where $\lambda(\cdot)$ satisfies (12). The proof is completed. □

Theorem 3 *Suppose that $[\mathcal{W}, \mathcal{R}_n]$ be a KB. Let $\boldsymbol{M}(\mathcal{W})$ be the measure set on $\mathcal{W}$, where $\mathcal{W} = \{w_i\}_k$, which can be divided by relation $\mathcal{R}_n = \{\boldsymbol{R}_j\}_n$. Then the $C_v(\boldsymbol{M}_{KEN}(\mathcal{W}))$ can be equivalently described by the measurement function $\lambda(x)$ (i.e., (12)).*

Proof Suppose that $[\mathcal{W}, \mathcal{R}_n]$ be a KB, and let $\mathbf{M}_{\mathrm{REN}}(\mathcal{W})$ be the measure set on the $\mathcal{W}$ based on rough entropy, we suppose that,

$$\begin{aligned} \mathbf{M}_{\mathrm{KEN}}(\mathcal{W}) &= \{\mathrm{KEN}(\mathcal{R}_1), \mathrm{KEN}(\mathcal{R}_2), ..., \mathrm{KEN}(\mathcal{R}_n)\} \\ &= \{\mathrm{KEN}(\mathcal{R}_j)\}_n. \end{aligned} \tag{27}$$

According to (11), then we obtain the following, i.e.,

$$\begin{aligned}\overline{\mathrm{KEN}(\mathcal{R})} &= \frac{1}{n}\sum_{j=1}^{n}\mathrm{KEN}(\mathcal{R}_j),\\ \sigma(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W})) &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KEN}(\mathcal{R}_j)-\overline{\mathrm{KEN}(\mathcal{R})}\right)^2},\\ C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W})) &= \frac{\sigma(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))}{\overline{\mathrm{KEN}(\mathcal{R})}}\end{aligned} \tag{28}$$

According to (5), for the set $\mathcal{W}=\{w_i\}_k$ (i.e., $|W|=k$), it follows that,

$$\mathrm{KEN}(\mathcal{R}_j)=-\sum_{i=1}^{k}\frac{1}{k}\log_2\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}=\frac{1}{k}\sum_{i=1}^{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|} \tag{29}$$

and

$$\overline{\mathrm{KEN}(\mathcal{R})}=\frac{1}{n}\sum_{j=1}^{n}\mathrm{KEN}(\mathcal{R}_j)=\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}. \tag{30}$$

Further, we can obtain

$$\begin{aligned}&\sigma(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))\\ &=\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KEN}(\mathcal{R}_j)-\overline{\mathrm{KEN}(\mathcal{R})}\right)^2}\\ &=\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{k}\frac{1}{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}-\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}\right)^2}\end{aligned} \tag{31}$$

, and

$$\begin{aligned}&C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))\\ &=\frac{\sigma(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))}{\overline{\mathrm{KEN}(\mathcal{R})}}\\ &=\frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KEN}(\mathcal{R}_j)-\overline{\mathrm{KEN}(\mathcal{R})}\right)^2}}{\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}}\\ &=\frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{k}\frac{1}{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}-\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}\right)^2}}{\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}}\\ &=\frac{\sqrt{n\cdot\sum_{i=1}^{n}\left(\sum_{i=1}^{k}(\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|})-\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{k}(\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|})\right)^2}}{\sum_{j=1}^{n}\sum_{i=1}^{k}(\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|})}.\end{aligned} \tag{32}$$

According to (32), we establish the mapping relationship between $C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))$ and $\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}$, i.e.,

$$\begin{aligned}\lambda\left(\log_2\left(\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}\right)\right) &= C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))\\ &\triangleq C_v(\mathrm{KEN}(\{\mathcal{R}_j\})_n),\end{aligned} \tag{33}$$

where $\lambda(\cdot)$ satisfies (12). The proof is completed. □

Theorem 4 *Suppose that $[\mathcal{W},\mathcal{R}_n]$ be a KB. Let $\boldsymbol{M}(\mathcal{W})$ be the measure set on $\mathcal{W}$, where $\mathcal{W}=\{w_i\}_k$, which can be divided by relation $\mathcal{R}_n=\{\boldsymbol{R}_j\}_n$. Then the $C_v(\boldsymbol{M}_{KAM}(\mathcal{W}))$ can be equivalently described by the measurement function $\lambda(x)$ (i.e., (12)).*

Proof Suppose that $[\mathcal{W},\mathcal{R}_n]$ be a KB, and let $\mathbf{M}_{\mathrm{RAM}}(\mathcal{W})$ be the measure set on the $\mathcal{W}$ based on rough entropy, we suppose that,

$$\begin{aligned}\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}) &= \{\mathrm{KAM}(\mathcal{R}_1),\mathrm{KAM}(\mathcal{R}_2),...,\mathrm{KAM}(\mathcal{R}_n)\}\\ &= \{\mathrm{KAM}(\mathcal{R}_j)\}_n.\end{aligned} \tag{34}$$

According to (11), then we obtain the following, i.e.,

$$\begin{aligned}\overline{\mathrm{KAM}(\mathcal{R})} &= \frac{1}{n}\sum_{j=1}^{n}\mathrm{KAM}(\mathcal{R}_j),\\ \sigma(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W})) &= \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KEN}(\mathcal{R}_j)-\overline{\mathrm{KAM}(\mathcal{R})}\right)^2},\\ C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W})) &= \frac{\sigma(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))}{\overline{\mathrm{KAM}(\mathcal{R})}}\end{aligned} \tag{35}$$

According to (7), for the set $\mathcal{W}=\{w_i\}_k$ (i.e., $|W|=k$), it follows that,

$$\mathrm{KAM}(\mathcal{R}_j)=\sum_{i=1}^{k}\frac{1}{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right), \tag{36}$$

and

$$\overline{\mathrm{KAM}(\mathcal{R})}=\frac{1}{n}\sum_{j=1}^{n}\mathrm{KAM}(\mathcal{R}_j)=\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right). \tag{37}$$

Further, we obtain that,

$$\begin{aligned}&\sigma(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))\\ &=\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KAM}(\mathcal{R}_j)-\overline{\mathrm{KAM}(\mathcal{R})}\right)^2}\\ &=\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{k}\frac{1}{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)-\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)\right)^2}\end{aligned} \tag{38}$$

and

$$
\begin{aligned}
&C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))\\
&=\frac{\sigma(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))}{\overline{\mathrm{KAM}(\mathcal{R})}}\\
&=\frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\mathrm{KAM}(\mathcal{R}_j)-\overline{\mathrm{KAM}(\mathcal{R})}\right)^2}}{\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)}\\
&=\frac{\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{k}\frac{1}{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)-\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)\right)^2}}{\frac{1}{nk}\sum_{j=1}^{n}\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)}\\
&=\frac{\sqrt{n\cdot\sum_{i=1}^{n}\left(\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)-\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)\right)^2}}{\sum_{j=1}^{n}\sum_{i=1}^{k}\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)}.
\end{aligned}
\tag{39}
$$

Therefore, we establish the mapping relationship between $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))$ and $\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)$, i.e.,

$$
\begin{aligned}
\lambda\left(1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\right)&=C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))\\
&\triangleq C_v(\mathrm{KAM}(\{\mathcal{R}_j)\}_n),
\end{aligned}
\tag{40}
$$

where $\lambda(\cdot)$ satisfies (12). The proof is completed. □

6.1 The relation between $\lambda(\cdot)$ and KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$) and KAM($\mathcal{R}$)

According to Theorems 1-4, we summarize the intrinsic properties of function $\lambda(\cdot)$. Specifically, we can capture the following three important pieces of information:

1. *Universality* Measurement function $\lambda(\cdot)$ establishes an internal relationship with $C_v(\cdot)$ (e.g., (19)), in the final mathematical expression, we find that the set $\mathcal{W}$ does not affect (12). In other words, (12) is applied to any finite set (only requires $\mathcal{W}$ can be divided according to some relation $\mathcal{R}$), which means that the function $\lambda(\cdot)$ has universality.
2. *One-to-one correspondence between four measurement functions and the input of* $\lambda(\cdot)$ For example, $\log_2|[w_i]_{\mathcal{R}_j}|$ corresponds to $\mathrm{REN}(\mathcal{R}_n)$; $1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}$ corresponds to $\mathrm{KAM}(\mathcal{R}_n)$. Therefore, $\lambda(\cdot)$ achieves formal unification of the four different measurement functions.
3. *Monotonicity* The function $\lambda(\cdot)$ can uniformly describe these four different measurement tools in a two-dimensional plane. Since $\left|[w_i]_{\mathcal{R}_j}\right| \in \mathbb{R}$, thus that, $|[w_i]_{\mathcal{R}_j}|$, $\log_2|[w_i]_{\mathcal{R}_j}|$, $\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}$ and $1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}$ can be described by the parameters x, $\log_2(x)$, $\log_2(\frac{k}{x})$, and $1-\frac{x}{k}$, where $x>0, k\in\mathbb{Z}^+$, and they are all elementary functions in a two-dimensional plane.

Equivalent representation According to $\lambda(\cdot)$ and $C_v(\cdot)$, we use $\lambda(\cdot)$ to describe $C_v(\cdot)$ equivalently. In addition, according to (12), we see that the difference between $C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))$, $C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}))$, $C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))$ and $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))$ are completely dependent on their different inputs $|[w_i]_{\mathcal{R}_j}|$, $\log_2|[w_i]_{\mathcal{R}_j}|$, $\log_2\frac{k}{\left|[w_i]_{\mathcal{R}_j}\right|}$ and $1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}$. Therefore, the difference between four mathematical tools for measuring the uncertainty of $[\mathcal{W},\mathcal{R}]$ can be represented by x, $\log_2(x)$, $\log_2(\frac{k}{x})$, and $1-\frac{x}{k}$.

Interval range Observably, considering the monotonicity of each function, we can obtain that in the interval $[\alpha,\beta]$, the In (10) always holds, where α satisfies $\alpha=x_1=\sqrt{k}$ (i.e., $\log_2(x_1)=\log_2(\frac{k}{x_1})$), and β satisfies $\beta=x_2=2k$ or $x_2=k$ (i.e., $1-\frac{x_2}{k}=\log_2(\frac{k}{x_2})$). Consequently, we obtain an initial range, that is $[\sqrt{k},2k]$, $k\in\mathbb{Z}^+$. However, $1-\frac{x_2}{k}=1-\frac{2k}{k}=-1$, which contradicts with $1-\frac{\left|[w_i]_{\mathcal{R}_j}\right|}{k}\geq 0$ (because $\left|[w_i]_{\mathcal{R}_j}\right|\leq k$). Then the value of β_{min} should be subject to $1-\frac{x_2}{k}=0$, i.e., $\beta=x_2'=k$. Therefore, we obtain that,

Corollary 1 *If*

$$
\left|[w_i]_{\boldsymbol{R}_j}\right|\in[\alpha,\beta]=[x_1,x_2']=[\sqrt{k},k]\subseteq[\lceil\sqrt{k}\rceil,k],
\tag{41}
$$

where $\lceil k\rceil$ *is ceiling function, i.e.,* $\lceil k\rceil=\min\{n\in\mathbb{Z}|k\leqslant n\}$ *(e.g.,* $\lceil 2.4\rceil=3$*). Then,*

$$
\begin{aligned}
C_v(\boldsymbol{M}_{KGR}(\mathcal{W}))&>C_v(\boldsymbol{M}_{REN}(\mathcal{W}))>C_v(\boldsymbol{M}_{KEN}(\mathcal{W}))\\
&>C_v(\boldsymbol{M}_{KAM}(\mathcal{W}))
\end{aligned}
$$

For an intuitive experience, we provide two visualizations of the different evaluation functions of x, $\log_2(x)$, $\log_2(\frac{k}{x})$, and $1-\frac{x}{k}$ under different k values. According to Fig. 1 ($k=16$), and Fig. 2 ($k=25$), we can clearly see the difference between the four measurement functions.

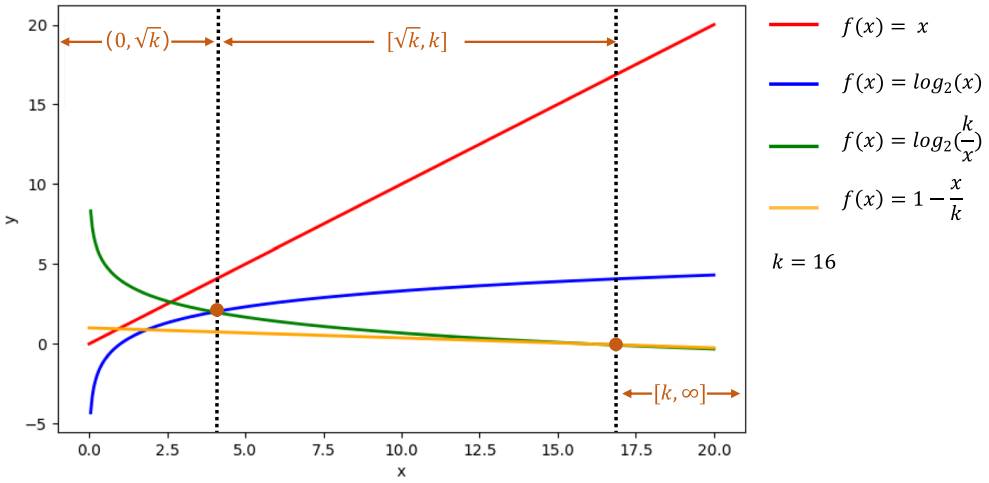

Fig. 1 A visualization of the different evaluation functions x, $\log_2(x)$, $\log_2(\frac{k}{x})$, and $1 - \frac{x}{k}$ at $k = 16$

Note We provide two visual examples to understand the unified representation of these four measurement functions, which correspond to the four different inputs of the unified metric function $\lambda(\cdot)$. In the previous section, we provide an explicit interval within which the Inequality (10) holds strictly. However, as shown in Figs. 1 and 2, the magnitude relations of the four measurement functions are not unique, if $\left|[w_i]_{\mathbf{R}_j}\right| \in (0, \sqrt{k})$. In summary, we conclude the following:

1. When $\left|[w_i]_{\mathbf{R}_j}\right| \in [\sqrt{k}, k]$, inequality (10) holds strictly. In other words, KAM($\mathcal{R}$) has a much better performance for measuring the uncertainty of KBs.
2. When $\left|[w_i]_{\mathbf{R}_j}\right| \in (0, \sqrt{k})$, the four measurement functions do not show regularity in the results, and KAM($\mathcal{R}$) almost always shows better performance. Note that since k represents the number of samples in the dataset, the interval $\left|[w_i]_{\mathbf{R}_j}\right| \in (k, +\infty)$ does not exist in practice, so we will not discuss this situation.

Comparison analysis $\lambda(\cdot)$ formally unifies KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$), and KAM($\mathcal{R}$). Next, we visualize the similarities and differences between $\lambda(\cdot)$ and KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$), and KAM($\mathcal{R}$) by Figs. 3 and 4.

It is worth noting that $\lambda(\cdot)$ is not a new measurement function, which is used as a unified equivalent form of KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$), and KAM($\mathcal{R}$). Therefore, the following analysis does not involve a comparison of performance, while focusing on the differences between $\lambda(\cdot)$ and each measurement function in terms of principle, interpretability. Specifically, as shown in Figs. 3 and 4, we summarize the comparison between $\lambda(\cdot)$ and KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$), and KAM($\mathcal{R}$) as follows:

1. **Measurement principle:** For KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$), and KAM($\mathcal{R}$), they focus only on outputting specific numerical results (e.g., coefficients of variation) in their studies of measures of uncertainty for knowledge bases. In other words, the comparison of the performance between these measurement functions

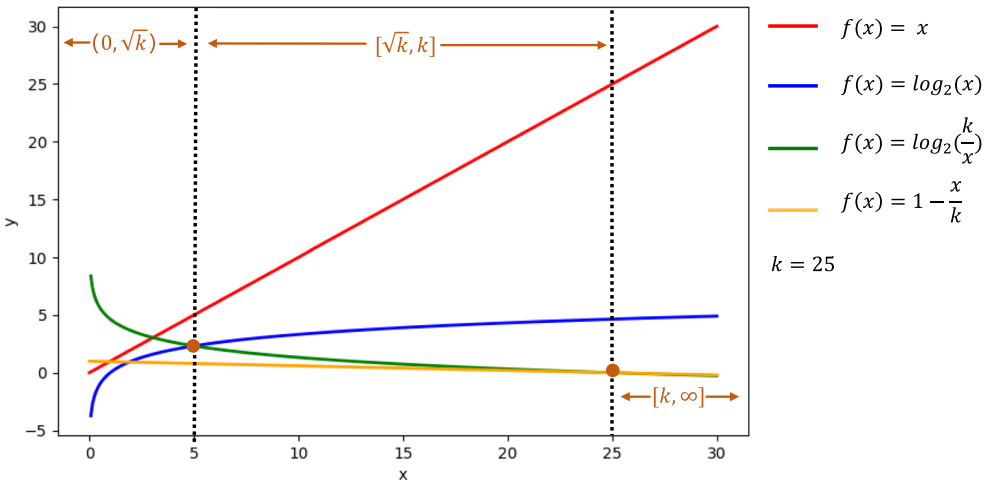

Fig. 2 A visualization of the different evaluation functions x, $\log_2(x)$, $\log_2(\frac{k}{x})$, and $1 - \frac{x}{k}$ at $k = 25$

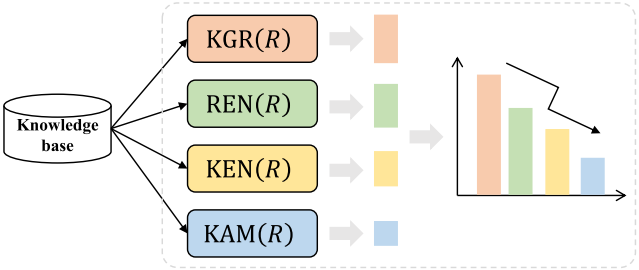

Fig. 3 Comparison of the measure values of the four measurement functions

is also limited to the presentation by the magnitude of the statistical values they compute. Unfortunately, this comparison at the level of results alone does not reflect why the four measurement functions differ. For example, in the case where the potential association between KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$), and KAM($\mathcal{R}$) are not considered, it does not reveal the reason, although it can reflect that the value of "pink" is (almost always) greater than the value of "blue" (as shown on the left in Fig. 3).

2. **Interpretability:** As shown in Fig. 4, $\lambda(\cdot)$ integrates the four measurement functions in a unified measurement framework, where different inputs correspond to different outputs. In Theorem 1, we have proved that $\lambda(\cdot)$ has the following form, i.e.,

$$\lambda(\cdot) = \frac{\sqrt{n \cdot \sum_{i=1}^{n} \left(\sum_{i=1}^{k}(\cdot) - \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{k}(\cdot) \right)^2}}{\sum_{j=1}^{n} \sum_{i=1}^{k}(\cdot)},$$
$$x = |[w_i]_{\mathcal{R}_j}| \in \mathbb{Z}^+.$$

Obviously, for determined x, n, and k (which can be determined from the knowledge base), $\lambda(\cdot)$ involves only changes in values and therefore does not change the monotonicity of the original input. This excellent property allows the comparison between different outputs based on $\lambda(\cdot)$ to be translated into a comparison of their corresponding inputs, i.e., x, $\log_2(x)$, $\log_2(\frac{k}{x})$, and $1 - \frac{x}{k}$. Fortunately, each of the above four inputs corresponds to four more primitive functions and can be compared (as shown in Figs. 1 and 2). Thus, although $\lambda(\cdot)$ is not a new measurement function, as a unified integrated framework for KGR($\mathcal{R}$), REN($\mathcal{R}$), KEN($\mathcal{R}$), and KAM($\mathcal{R}$), it explains the differences in the metric values of different measurement functions by comparing x, $\log_2(x)$, $\log_2(\frac{k}{x})$, and $1 - \frac{x}{k}$.

Limitations In RST, knowledge reflects the ability to classify some objects [45]. Specifically, in a KB, the set of entities we are interested in a certain field can be regarded as a finite set (or *universe*) $\mathcal{W}$ and any subset $\mathcal{C} \subseteq \mathcal{W}$ is called a *category* (or *concept*) in $\mathcal{W}$, which contains many entities. The *concept family*, which contains many concepts, is called *abstract knowledge* about $\mathcal{W}$. A KB over $\mathcal{W}$ is equivalent to a family of classifications over $\mathcal{W}$. Objects in a KB can be divided according to their different attributes. For example, given a set $\mathcal{W}$, which contains many `candies`, and suppose these candies have different `colors` (e.g., white, yellow, red) and `shapes` (e.g., round, square, triangle), then, these candies can be described by attributes such as `color` and `shape`, e.g., `red round candies`, or `yellow triangle candies`, etc. According to different attributes, we can describe the specific situation of these candies by a certain attribute (e.g., `color` and `shape`). Hence, we can obtain two equivalence relations (or attributes) from the above example, i.e., $\mathcal{R} = \{\mathbf{R}_1, \mathbf{R}_2\} = \{\texttt{color}, \texttt{shape}\}$. According to these equivalence relations, the corresponding equivalence class can be further obtained. The elements in the set $\mathcal{W}$ are divided and recombined according to the equivalence relations, e.g., `candies` are divided by `color`.

7 Measures of uncertainty for KBs without attribute information

In the previous section, we analyze the performance of different measurement functions in measuring the uncertainty of KBs. The limitation of previous research is that the division of instances in a KB can often only depend on their attributes. However, the type of knowledge base

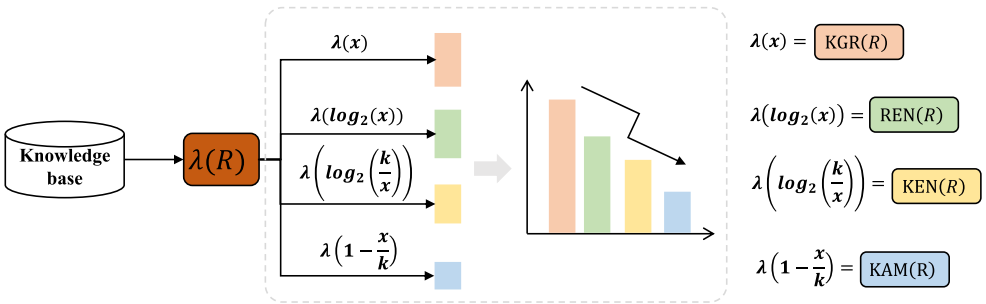

Fig. 4 Comparison of the outputs in $\lambda(\cdot)$ corresponding to the four different inputs

has changed with the needs of real applications, and some of the knowledge bases do not contain the attributes of the instances or lack sufficient attribute relations to classify the instances (e.g., ProBase). In this section, we first provide the definition of *concept structure* of ProBase (see Definition 6). And then, we provide an effective strategy to induce KBs from ProBase, and instances in induced KBs can be classified by their concepts.

7.1 Inducing KBs from ProBase: intuition

According to the definition 4, for the sake of simplicity of description, we use a $[\mathcal{T}, \mathcal{H}]$ to represent a KB induced by ProBase. In fact, in ProBase, all KBs are induced by the same strategy. Hence, in the rest of this paper, we unify all knowledge bases into $[\mathcal{T}, \mathcal{H}]$ for theoretical analysis. Specifically, the more accurate description is that $\mathcal{T}$ is the set containing a large number of *instances*, which refer to nodes that no longer have hyponyms in Pobase, and $\mathcal{H}$ is the family of *hypernyms* (or *concepts*) set of instances. Therefore, in this paper, we do not strictly distinguish the difference between *InstanceOf* and *SubClass*. In most downstream tasks, the two can be unified as the *isA* relationship.

Definition 12 (ProBase [34]) ProBase[3] is probabilistic of taxonomy, which contains hundreds of millions of *instances*, *concepts*, and *isA* relationships. *isA* relationship can be specified as *InstanceOf* relation between a concept and an instance (e.g., (`Snoopy`, *isA*, `dog`)) or *SubClass* relation between a pair of concepts (e.g., (`fruit`, *isA*, `botany`)).

Classifications We first use a simple example to illustrate the intuition that the instances in ProBase can be classified according to their concepts.

Example 2 Given a finite set $\mathcal{T}_1 = \{\text{dhol, tiger, lion, wolf}\}$, if $\mathcal{T}_1$ is divided by the equivalence relation $\mathbf{H}_a = \{\text{carnivore}\}$, the equivalence class of $\mathcal{T}_1$ can form an independent set, i.e., $\mathcal{T} = \mathcal{C}$, where

$$\mathcal{C} = \mathcal{T}_1/\mathbf{H}_a = [\text{dhole, tiger, lion, wolf }]_{\mathbf{H}=\text{carnivore}}.$$

If $\mathcal{T}_1$ is divided by the equivalence relation

$$\mathbf{H}_b = \{\text{beast division}\} = \{\mathbf{H}_1, \mathbf{H}_2\} = \{\text{canidae, felidae}\}.$$

Then $\mathcal{T}_1$ can be divided into $\mathcal{C} = \mathcal{T}_1/\mathbf{H}_b = \{\mathcal{C}_1, \mathcal{C}_2\}$, where

$$\mathcal{C}_1 = \{\text{dhole, wolf}\}_{\mathbf{R}_1=\text{canidae}}, \text{ and } \mathcal{C}_2 = \{\text{tiger, lion}\}_{\mathbf{R}_2=\text{felidae}}.$$

As can be seen from Example 2, $\mathcal{T}_1$ can be divided by $\mathbf{H}_b \in \mathcal{H}$ to obtain $\mathcal{C}_1$ and $\mathcal{C}_2$.

[3] https://www.microsoft.com/en-us/research/project/probase/

For ProBase, the dimension of $\mathcal{T} = \{\mathcal{C}_i\}_m$ can be determined by $\mathcal{H}$, hence, $\mathcal{T} = \{\mathcal{C}_i\}_m$ can be regarded as a vector in vector space. Note that, suppose $[\mathcal{T}, \mathcal{H}]$ be a KB induced by ProBase, where $\mathcal{T}$ is the set of instances, and $\mathcal{H}$ is the family consisting of the set of hypernyms (i.e., concepts) of instances, then the choice of concepts is constrained. This means that the instances in $\mathcal{T}$ can be divided by $\mathcal{H}$. Therefore, in this paper, we regard an equivalence relation (i.e., attribute) in the KB as a concept (i.e., hypernym) in ProBase. Li et al. [33] define the vector $\mathcal{T} = \{\mathcal{C}_i\}_m$ as the knowledge structure of KBs. Similarly, we provide the definition of the *concept structures* of $[\mathcal{T}, \mathcal{H}]$ as follows:

Definition 13 (Concept structures of $[\mathcal{T}, \mathcal{H}]$) Suppose $[\mathcal{T}, \mathcal{H}]$ be a KB induced by ProBase, if the finite set $\mathcal{T} = \{t_i\}_k$ can be divided by relations $\mathcal{H} = \{\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_i\}$, then we call the vector

$$\text{CSV}(\mathcal{H}) = \left\langle [t_1]_{\mathcal{H}}, [t_2]_{\mathcal{H}}, \ldots, [t_k]_{\mathcal{H}} \right\rangle \tag{42}$$

is the concept structure of $[\mathcal{T}, \mathcal{H}]$.

In Example 2, let $t_1 =$ `tiger`, $t_2 =$ `lion`, and $\mathbf{H}_2 = \{$`felidae`$\}$, then $[t_1]_{\mathbf{H}_b} \triangleq [t_2]_{\mathbf{H}_b}$, which mean that `tiger` and `lion` are equivalent under relation $\mathbf{H}_2$. Similarity, $\mathcal{C}_1$ and $\mathcal{C}_2$ are equivalent under relation $\mathbf{H}_b$.

7.2 Inducing KBs from ProBase: strategy

Strategy It is worth noting that in ProBase, most instances belong to many hypernyms, in other words, two or more different concepts may have the identical instances (e.g., the hypernyms of `apple` can be `company`, `fruit`, etc.). Therefore, intuitively, ProBase can divide instances based on different levels of hypernyms to obtain multiple KBs, and the specific division strategy is:

1. Select an instance $t_i \in \mathcal{T}$ which should have at least three hypernym hierarchies (denoted as $h^j(t_i, q), i \in |\mathcal{T}|, j, q \in \mathbb{Z}^+, j_{max} \geqslant 3$), i.e.,

$$t_i \longrightarrow h_k^1(t_i, q) \longrightarrow h^2(t_i, q) \longrightarrow h^3(t_i, q) \longrightarrow \cdots, \tag{43}$$

where $x \longrightarrow y$ means x is the hyponym of y. For example,

$$\texttt{corn} \longrightarrow \texttt{crop} \longrightarrow \texttt{plant} \longrightarrow \cdots \tag{44}$$

2. Repeat the above strategy, and finally obtain all $h_k^1(t_i)$ satisfying (45), i.e.,

$$t_i \longrightarrow \left\{ \begin{array}{c} h_1^1(t_i, 1) \\ h_2^1(t_i, 1) \\ \vdots \\ h_k^1(t_i, 1) \end{array} \right\} \longrightarrow h^2(t_i, 1) \longrightarrow \cdots \tag{45}$$

For example.

$$\begin{aligned}&\texttt{corn} \longrightarrow \texttt{crop} \longrightarrow \texttt{plant} \longrightarrow \cdots,\\ &\texttt{corn} \longrightarrow \texttt{Monocotyledoneae} \longrightarrow \texttt{plant} \longrightarrow \cdots,\\ &\texttt{corn} \longrightarrow \texttt{herbaceous plants} \longrightarrow \texttt{plant} \longrightarrow \cdots.\end{aligned} \tag{46}$$

$$\texttt{corn} \longrightarrow \begin{cases} \texttt{food} \longrightarrow \texttt{Foods Association} \longrightarrow \cdots, \\ \vdots \\ \texttt{coarse food grain} \longrightarrow \texttt{Foods Association} \longrightarrow \cdots. \end{cases} \tag{47}$$

5. Until t_i does not satisfy (45), the search is terminated. The final acquired dataset

$$\begin{aligned}&[\mathcal{T}, \mathcal{H}],\\ &\mathcal{T} = \{T_1, T_2, ..., T_q\},\\ &\mathcal{H} = \{h^2(t_i, 1), h^2(t_i, 2), ..., h^2(t_i, q)\}.\\ &\text{s.t.} \begin{cases} T_i \cap T_{j,j\neq i} = \emptyset, \\ hypo(h^2(t_i, q_i)) \cap hypo(h^2(t_i, q_{j,j\neq i})) \neq \emptyset. \end{cases}\end{aligned} \tag{48}$$

can be viewed as a sub-dataset induced by ProBase, based on instance t_i. $T_i \cap T_{j,j\neq i} = \emptyset$ ensures that the same *instance* is strictly divided according to its *hypernyms*. For example, a `candy` cannot be both `red` and `blue`. $hypo(h^2(t_i, q_i)) \cap hypo(h^2(t_i, q_{j,j\neq i})) \neq \emptyset$ ensures that presence of instances under any combination of $hypo(h^2(t_i, q_i), q_i \in \{1, 2, , ..., q\})$.

Rationality analysis The strategy is not unique. Similarly, we also select a concept (the concept must have enough hypernym hierarchies and hyponym hierarchies) to conform to the selection strategy of (45). We won't repeat it here. Obviously, multiple KBs can be induced from ProBase based on the above strategy, and the instances in these KBs can be divided according to their selected concepts. As a comparison, in $[\mathcal{T}, \mathcal{H}]$, a "$h^2(t_i, q)$" plays the role of an *attribute*, and "$h^1_k(t_i, 1)$" represents the *attribute value*. Therefore, based on the above strategy and analysis, we theoretically provide a strategy for inducing a KB from ProBase, and the instances in the induced KB can be strictly classified based on their selected concepts. Our results indicate that $\lambda(\cdot)$ provides valuable insights to integrate four measurement functions into a unified framework for measuring the uncertainty of KBs.

8 Experiments

8.1 KBs with attribute information

Comparison of four measurement functions We conduct experiments on the datasets in Table 4 with the aim of comparing the performance of four measurement functions, KGR$(\cdot)$, REN$(\cdot)$, KEN$(\cdot)$ and KAM$(\cdot)$, across different knowledge bases.

3. Collect all the instances in each $h^1_k(t_i, 1)$ to form set T_1.
4. Repeat the selection strategy above, similarly, we collect all the instances in each $h^1_{k'}(t_i, 2)$ to form set T_2.

For example,

The measure sets construction Specifically, for a KB $[\mathcal{W}, \mathcal{R}]$, we denote $R_i = ind(\{f_i \in \mathcal{R}\})$, where $ind(\cdot)$ stands for the indiscernibility relation, such as $ind(\mathcal{R}) = \bigcap_{f_i \in \mathcal{R}} \mathcal{R}$. Let $\mathcal{R}$ be the set consisting of R_i, where $\mathcal{R}$ satisfies $\mathcal{R}_j = \{R_1, R_2, ..., R_j\}$ (e.g., $\mathcal{R}_3 = \{R_1, R_2, R_3\}$). Obviously, $[\mathcal{W}, \mathcal{R}_j]$ is the knowledge base induced by $\mathcal{W}$. Therefore, we obtain four measure sets on $\mathcal{W}$ as follows:

$$\begin{aligned}&\mathbf{M}(\text{KGR}) = \{\text{KGR}(\mathcal{R}_1), \text{KGR}(\mathcal{R}_2), ..., \text{KGR}(\mathcal{R}_j)\},\\ &\mathbf{M}(\text{REN}) = \{\text{REN}(\mathcal{R}_1), \text{REN}(\mathcal{R}_2), ..., \text{REN}(\mathcal{R}_j)\},\\ &\mathbf{M}(\text{KEN}) = \{\text{KEN}(\mathcal{R}_1), \text{KEN}(\mathcal{R}_2), ..., \text{KEN}(\mathcal{R}_j)\},\\ &\mathbf{M}(\text{KAM}) = \{\text{KAM}(\mathcal{R}_1), \text{KAM}(\mathcal{R}_2), ..., \text{KAM}(\mathcal{R}_j)\},\end{aligned} \tag{49}$$

Example 3 For example, "Lymphography" in Table 4 can be viewed as an information system $[\mathcal{T}, \mathcal{F}]$ with $|\mathcal{T}| = 148$, $|\mathcal{F}| = 18$. We can obtain four measure sets on "Lymphography" as follows:

$$\begin{aligned}&\mathbf{M}_{\text{KGR}}(\mathcal{W}) = \{\text{KGR}(\mathcal{R}_1), \text{KGR}(\mathcal{R}_2), ..., \text{KGR}(\mathcal{R}_{18})\},\\ &\mathbf{M}_{\text{REN}}(\mathcal{W}) = \{\text{REN}(\mathcal{R}_1), \text{REN}(\mathcal{R}_2), ..., \text{REN}(\mathcal{R}_{18})\},\\ &\mathbf{M}_{\text{KEN}}(\mathcal{W}) = \{\text{KEN}(\mathcal{R}_1), \text{KEN}(\mathcal{R}_2), ..., \text{KEN}(\mathcal{R}_{18})\},\\ &\mathbf{M}_{\text{KAM}}(\mathcal{W}) = \{\text{KAM}(\mathcal{R}_1), \text{KAM}(\mathcal{R}_2), ..., \text{KAM}(\mathcal{R}_{18})\},\end{aligned} \tag{50}$$

and the values of KGR$(\mathcal{R}_j)$, REN$(\mathcal{R}_j)$, KEN$(\mathcal{R}_j)$ and KAM$(\mathcal{R}_j)$ are calculated by (4)–(7).

8.2 Experimental results and analysis on multi-domain datasets

Experimental results The experimental results are shown in Table 5 and Fig. 5.

Analysis From the results, we conclude that:

1. **Consistency of results:** We select datasets from different domains to validate our theoretical analysis, which contains different numbers of instances and attributes. Specifically, 18 datasets involving 6 domains (i.e.,

Table 4 Data sets from UCI,[a] "#X"represents the number of "X"

Datasets	Area	#Attributes	#Instances
Tic-Tac-Toe Endgame	Game	9	958
Chess	Game	36	3,196
Dota2 Games	Game	116	102,944
Lymphography	Life Science	18	148
Mushroom	Life Science	22	8,124
SPECT Heart	Life Science	22	267
Abalone	Life Science	8	4,177
Estimation of obesity levels	Life Science	17	2,111
Primary Tumor	Life Science	17	339
Breast Cancer	Life Science	10	116
Congressional Voting Records	Social Science	16	435
Balance Scale	Social Science	4	625
Nursery	Social Science	8	12,960
Student Performance	Social Science	33	649
Letter Recognition	Computer	16	20,000
Solar Flare	Physical	10	1,389
Car Evaluation	Other	6	1,728
MONK's Problems	Other	7	432

[a] https://archive.ics.uci.edu/ml/index.php

game, life science, social science, computer, physical and other) all consistently demonstrate our theoretical analysis, i.e.,

$$C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W})).$$

2. **Metric Performance:** For the dataset of different domains, the value of $C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))$ fluctuates the most, and it has the worst performance for measuring the uncertainty of KBs. By contrast, the value of $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))$ has good stability, and it has the best performance for measuring the uncertainty of KBs.

Table 5 Coefficient of variation values of measure sets $\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}_i)$, $\mathbf{M}_{\mathrm{REN}}(\mathcal{W}_i)$, $\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}_i)$, and $\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}_i)$

Index	Datesets	$C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}_i))$	$C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}_i))$	$C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}_i))$	$C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}_i))$
$\mathcal{W}_1$	Tic-Tac-Toe Endgame	1.7879	0.9015	0.4340	0.1186
$\mathcal{W}_2$	Chess	1.5765	0.6719	0.5865	0.1276
$\mathcal{W}_3$	Dota2 Games	4.6868	2.5229	1.5775	0.6037
$\mathcal{W}_4$	Lymphography	1.5971	0.7135	0.4518	0.0946
$\mathcal{W}_5$	Mushroom	2.8592	0.6501	0.3279	0.0807
$\mathcal{W}_6$	SPECT Heart	0.8096	0.5593	0.2969	0.1384
$\mathcal{W}_7$	Abalone	2.0676	1.7854	0.6837	0.1041
$\mathcal{W}_8$	Estimation of obesity levels	3.6314	3.0076	0.2442	0.1288
$\mathcal{W}_9$	Primary Tumor	1.8870	0.8839	0.3288	0.1289
$\mathcal{W}_{10}$	Breast Cancer	1.5247	0.9560	0.3517	0.0980
$\mathcal{W}_{11}$	Congressional Voting Records	1.5189	0.6481	0.3574	0.1253
$\mathcal{W}_{12}$	Balance Scale	1.2943	0.7453	0.4472	0.0861
$\mathcal{W}_{13}$	Nursery	2.0431	0.6978	0.4750	0.1141
$\mathcal{W}_{14}$	Student Performance	3.1088	1.9325	0.2946	0.1643
$\mathcal{W}_{15}$	Letter Recognition	3.1032	1.3883	0.2953	0.0380
$\mathcal{W}_{16}$	Solar Flare	0.9537	0.4224	0.1988	0.0204
$\mathcal{W}_{17}$	Car Evaluation	1.5439	1.0686	0.2148	0.0556
$\mathcal{W}_{18}$	MONK's Problems	1.3650	0.9847	0.3916	0.1201

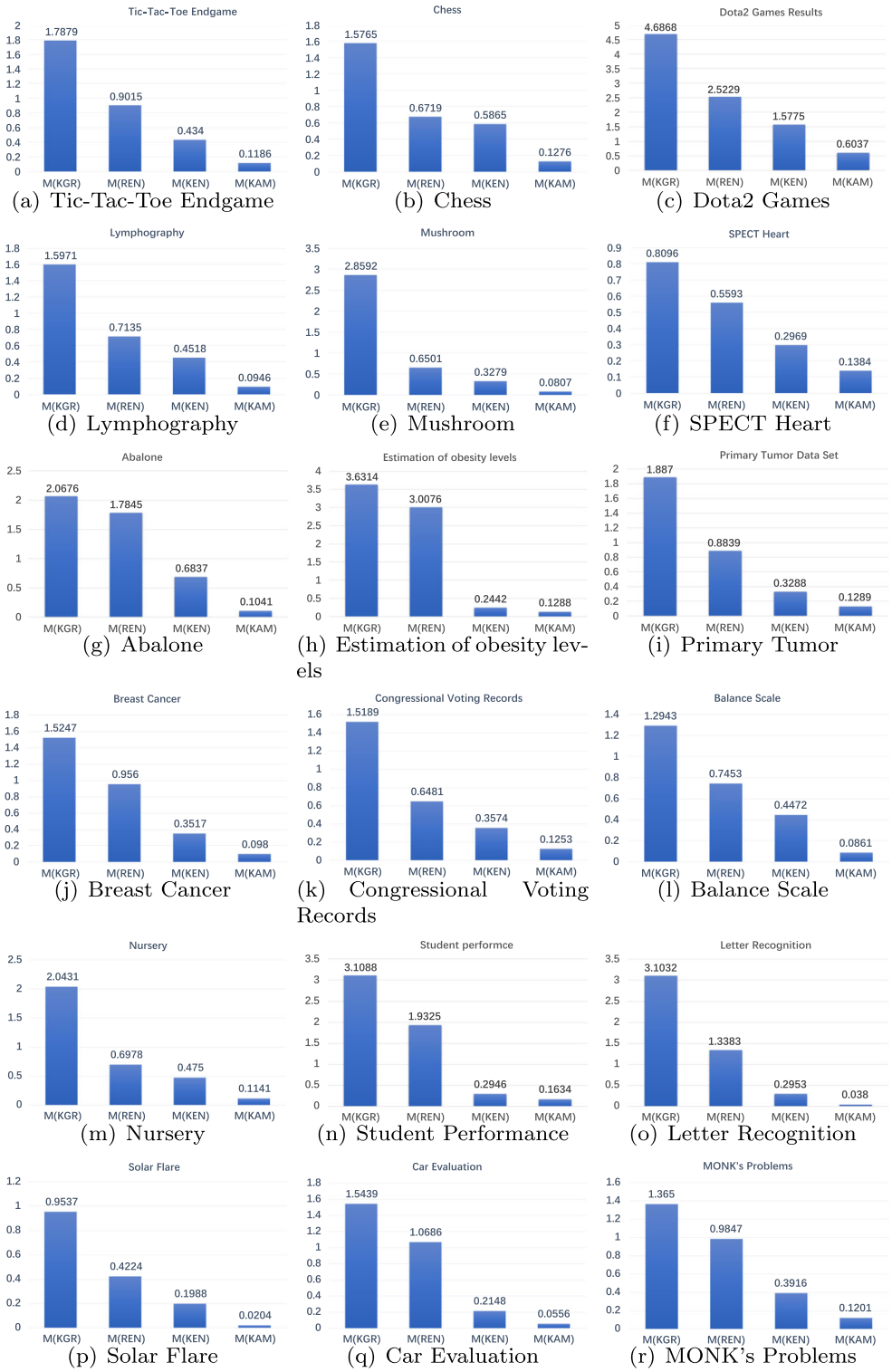

Fig. 5 Coefficient of variation values of four measure sets on datasets (**a**)–(**r**)

3. **Comparison of $C_v(\mathbf{M_{REN}}(\mathcal{W}))$ and $C_v(\mathbf{M_{KEN}}(\mathcal{W}))$:** As shown in Fig. 5, the gap between $C_v(\mathbf{M_{REN}}(\mathcal{W}))$ and $C_v(\mathbf{M_{KEN}}(\mathcal{W}))$ is not significant in most of the datasets, which is consistent with our analysis of the measurement functions REN($\mathcal{R}$) and REN($\mathcal{R}$) in the previous section. For example, as shown in Figs. 1 and 2, when the value of x is in the interval $[\sqrt{k}, k]$, the gap between $C_v(\mathbf{M_{REN}}(\mathcal{W}))$ and $C_v(\mathbf{M_{KEN}}(\mathcal{W}))$ is not too significant in most cases.
4. **Comparison of $C_v(\mathbf{M_{RGR}}(\mathcal{W}))$ and $C_v(\mathbf{M_{KAM}}(\mathcal{W}))$:** Contrasted with the above conclusion, the gap between $C_v(\mathbf{M_{RGR}}(\mathcal{W}))$ and $C_v(\mathbf{M_{KAM}}(\mathcal{W}))$ demonstrates a

Table 6 Statistical information of D_1, D_2 and D_3

Datasets	#concepts ($h^2(t_i, q)$)	#Instances
D_1	3	72
D_2	3	123
D_3	3	1290

significant difference on almost all datasets, which is consistent with our analysis of the measurement functions RGR($\mathcal{R}$) and KAM($\mathcal{R}$) in the previous section. For example, as shown in Figs. 1 and 2, when the value of x is in the interval $[\sqrt{k}, k]$, the gap between $C_v(\mathbf{M}_{\mathrm{RGR}}(\mathcal{W}))$ and $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))$ will increase as x increases.

8.3 KBs induced by ProBase

In this section, we aim to induce several KBs from ProBase based on the above strategy and to perform uncertainty measurement on the induced KBs. Specifically, we induce three different sizes of KBs (denoted as $D_1, D_2, and D_3$) for the metric, and the specific information of D_1 (based on concept `fruit` induction), D_2 (based on concept `corn` induction, containing 123 instances) and D_3 (based on concept `corn` induction, containing 1290 instances) are shown in Table 6. The construction method of the measure sets on D_1, D_2, and D_3 is the same as the construction method (49) on the general datasets.

8.4 Experimental results and analysis on ProBase

Experimental results The experimental results are shown in Table 7 and Fig. 6.

Analysis From the results, we conclude that:

1. In datasets D_1 and D_3, the results show the following relationship, i.e.,

$$C_v(\mathbf{M}_{\mathrm{KGR}}(D_i) > C_v(\mathbf{M}_{\mathrm{KEN}}(D_i)) > C_v(\mathbf{M}_{\mathrm{REN}}(D_i) > C_v(\mathbf{M}_{\mathrm{KAM}}(D_i). \quad (51)$$

The result is in line with our analysis conclusion. As shown in Figs. 1 and 2, we find that, in the interval $(0, \sqrt{k})$, there will be a situation where

$$C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})), \text{ if } \left|[w_i]_{\mathbf{R}_j}\right| \in [0, \sqrt{k}],\ w_i \in \mathcal{W}. \quad (52)$$

This fully validates the rigor of our theoretical analysis. Moreover, this conclusion also reveal that KEN($\mathcal{W}$) and REN($\mathcal{W}$) are greatly affected by the parameter k.

2. In dataset D_2, the results reveal the following relationship, i.e.,

$$C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W})).$$

This further verifies that $C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}))$ has stable and excellent performance in measuring the uncertainty of the KB.

3. Consistent with the experimental conclusions on the public datasets, KGR($\mathcal{W}$) has the worst performance in measuring the uncertainty of KBs, while KAM($\mathcal{W}$) maintains the best performance in measuring the uncertainty of KBs.

9 Case study

In this section, we provide a small-scale case to visually demonstrate how to use rough set theory and induction strategy (i.e., Section 7.2) to induce a measurable knowledge base (denoted as D_4) from ProBase. Dataset D_4 contains 19 concepts about `fruit`, and their corresponding hypernyms in ProBase (the selection of hypernyms is based on the induction strategy in Section 7.2). The statistical information of D_4 is summarized in Table 8.

Further, as in the above experiments, we construct measure sets on D_4, and calculate the coefficient of variation of measure sets, and the results are shown in Fig. 7.

Obviously, the experimental results based on dataset D_4 are consistent with the previous theoretical analysis and experimental evaluation conclusions. That is KGR(D_4) has the worst performance in measuring the uncertainty of KBs, while KAM(D_4) maintains the best performance in

Table 7 Coefficient of variation values of measure sets $\mathbf{M}_{\mathrm{KGR}}(D_i)$, $\mathbf{M}_{\mathrm{REN}}(D_i)$, $\mathbf{M}_{\mathrm{KEN}}(D_i)$, and $\mathbf{M}_{\mathrm{KAM}}(D_i)$ on dataset $D_{i,i=1,2,3}$

Datesets	$C_v(\mathbf{M}_{\mathrm{KGR}}(D_i))$	$C_v(\mathbf{M}_{\mathrm{REN}}(D_i))$	$C_v(\mathbf{M}_{\mathrm{KEN}}(D_i))$	$C_v(\mathbf{M}_{\mathrm{KAM}}(D_i))$
D_1	0.6217	0.3554	0.4246	0.2498
D_2	0.8889	0.5106	0.4073	0.1239
D_3	0.2705	0.0891	0.2397	0.0658

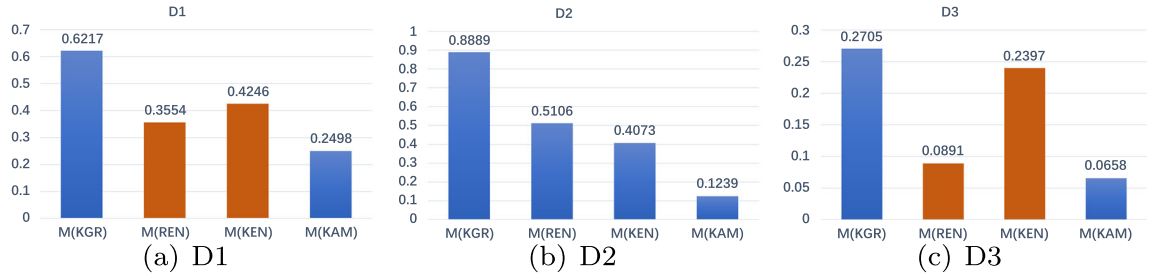

Fig. 6 Coefficient of variation values of four measure sets on datasets D_1, D_2 and D_3

measuring the uncertainty of the KB. In particular, the case study also captures the situation where $C_v(\mathbf{M}_{\mathrm{KEN}}(D_4))$ is greater than $C_v(\mathbf{M}_{\mathrm{REN}}(D_4))$.

10 Discussion

In this section, we hope to bring some guidance and insight to the study of knowledge base uncertainty through the results of the theoretical analysis in this paper. According to Table 5 and Fig. 5, we visually observe that although $C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))$, $C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}))$, $C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W}))$, and $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))$ exhibit the theoretical analysis of this paper on all 18 public datasets, i.e.,

$$C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{KEN}}(\mathcal{W})) > C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W})).$$

However, a more detailed analysis reveals that there are significant differences between the different measurement functions (e.g., in the dataset "Letter Precognition", $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))$ is 0.0380, but $C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))$ can reach 3.1032). Therefore, a single conclusion based on a single measurement function is not sufficient. Based on the theoretical analysis and experimental validation in this paper, we advocate that the uncertainty of the knowledge base should be evaluated by combining the four measurement functions. For example, for datasets"Solar Flare" and "Letter Recognition", although they differ slightly in the $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}))$ ($C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}_{15})) = 0.0380$, $C_v(\mathbf{M}_{\mathrm{KAM}}(\mathcal{W}_{16})) = 0.0204$), they differ significantly in the $C_v(\mathbf{M}_{\mathrm{KGR}}(\mathcal{W}))$ and $C_v(\mathbf{M}_{\mathrm{REN}}(\mathcal{W}))$. Therefore, it may be a more reasonable way to comprehensively consider these measurement functions.

Table 8 Statistical information of D_4

Fruits	Hard	Soft	Non-citrus	Citrus
Apple	✓		✓	
Apricot	✓		✓	
Banana	✓		✓	
Berry		✓	✓	
Cherry		✓		✓
Gooseberry		✓		✓
Grape		✓	✓	
Grapefruit		✓		✓
Kiwi		✓		✓
Melon		✓		✓
Orange	✓			✓
Papaya	✓			✓
Peach	✓		✓	
Pear	✓		✓	
Pineapple		✓	✓	
Plum	✓		✓	
Raspberry		✓		✓
Tomato		✓	✓	

The rapid development of deep neural networks (DNNs) in recent years has reached almost every field of AI, meanwhile, many researchers begin to think deeply about the reliability of prediction results based on neural networks. There is already evidence that uncertainty (e.g., data uncertainty and model uncertainty) imposes many limitations on DNNs, such as the lack of transparency of a DNN's inference framework [46]. In the previous sections, we focus on measures of uncertainty for knowledge bases, aiming to provide a rigorous theoretical analysis for the existing conclusions (e.g., uncover the reasons for performance differences between measurement functions). We hope these results will provide insights into understanding the essence of

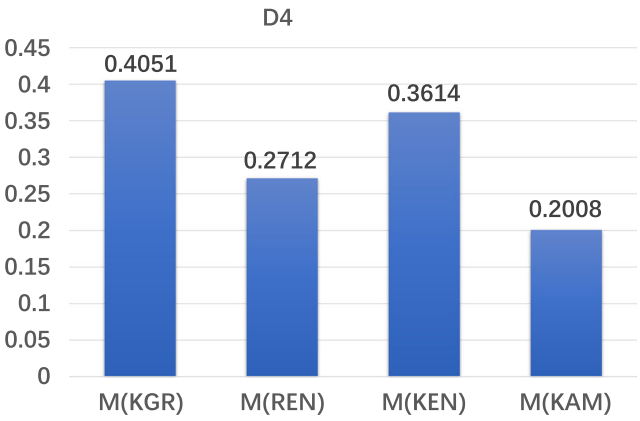

Fig. 7 Coefficient of variation values of measure sets on dataset D_4

uncertainty (e.g, uncertainty quantification [47]) for knowledge bases.

11 Conclusion and further work

The work of this paper is inspired by the experimental conclusions of [1]. In [1], the authors verify the superiority of measuring the uncertainty of KBs based on the knowledge amount through experiments on three datasets. Although this conclusion lacks rigorous theoretical analysis, it encourages us to study why the knowledge-amount-based measurement function has the best performance in measuring the uncertainty of the knowledge base. Therefore, this paper provides deeper insights into the uncertainty measurement of the knowledge base.

In this paper, we review four popular measurement functions in measuring the uncertainty for KBs. Then, at the theoretical level, we integrate the four measurement functions into a unified new measurement function, which provides valuable insights for measuring the uncertainty of KBs. At the experimental level, the experimental results on the 18 public datasets are consistent with our theoretical analysis conclusions, which fully demonstrates the correctness of our theoretical analysis. In addition, for some special datasets (e.g., ProBase), which contains a large amount of structured knowledge, there are not enough attributes to classify the instances in it. This leads to the inability of the above measurement functions to perform the uncertainty measurement on ProBase. In order to solve this issue, we propose an effective strategy, which can induce sub-datasets from ProBase, and all the instances in the sub-dataset can be divided according to their concepts. Comparative experimental results justify the effectiveness of the strategy and the consistency with the theoretical conclusions.

Further work Knowledge base, as an indispensable carrier for the development of artificial intelligence technology today, provides far-reaching resources for smart devices. With the increase in the amount of downstream real tasks and the diversification of real application scenarios, various types of knowledge bases have appeared one after another, and their knowledge structures have become more and more complicated. Therefore, how to measure the uncertainty of these knowledge bases is the future important work.

In addition, the timeliness, accuracy, and redundancy of the knowledge base are also important indicators to measure the knowledge base. Whether a complete theoretical analysis of the above measurement indicators can be established is one of our future efforts.

References

1. Li Z, Gangqiang Z, Wu W-Z, Xie N (2020) Measures of uncertainty for knowledge bases. Knowl Inf Syst 62(2):611–637
2. McDowell J, Brown L et al (2014) Theaetetus. Oxford University Press
3. Ferchichi A, Boulila W, Farah IR (2018) Reducing uncertainties in land cover change models using sensitivity analysis. Knowl Inf Syst 55(3):719–740
4. Resconi G, Kovalerchuk B (2009) Agents' model of uncertainty. Knowl Inf Syst 18(2):213–229
5. Eekhout JP, Millares-Valenzuela A, Martínez-Salvador A, GarcÍA-Lorenzo R, Pérez-Cutillas P, Conesa-García C, de Vente J (2021) A process-based soil erosion model ensemble to assess model uncertainty in climate-change impact assessments. Land Degrad Dev
6. Ghahramani Z (2015) Probabilistic machine learning and artificial intelligence. Nature 521(7553):452–459
7. Guo K, Xu H (2021) Preference and attitude in parameterized knowledge measure for decision making under uncertainty. Appl Intell, 1–10
8. Sun L, Guo J, Zhu Y (2019) Applying uncertainty theory into the restaurant recommender system based on sentiment analysis of online Chinese reviews. World Wide Web 22(1):83–100
9. Li R, Chen Z, Li H, Tang Y (2021) A new distance-based total uncertainty measure in Dempster-Shafer evidence theory. Appl Intell, 1–29
10. Wu Y, Lin X, Yang Y, He L (2019) Cleaning uncertain graphs via noisy crowdsourcing. World Wide Web 22(4):1523–1553
11. Zhu J, Ghosh S, Wu W (2020) Robust rumor blocking problem with uncertain rumor sources in social networks. World Wide Web, pp 1–19
12. Gambo S, Özad B (2021) The influence of uncertainty reduction strategy over social network sites preference. J Theor Appl Electron Commer Res 16(2):116–127
13. Ghasemi M, Bagherifard K, Parvin H, Nejatian S, Pho K-H (2021) Multi-objective whale optimization algorithm and multi-objective grey wolf optimizer for solving next release problem with developing fairness and uncertainty quality indicators. Appl Intell, pp 1–30
14. Kim J, Kim J, Wang Y (2021) Uncertainty risks and strategic reaction of restaurant firms amid covid-19: evidence from China. Int J Hosp Manag 92:102752
15. Albulescu CT (2021) Covid-19 and the United States financial markets' volatility. Finance Res Lett 38:101699
16. Viner RM, Bonell C, Drake L, Jourdan D, Davies N, Baltag V, Jerrim J, Proimos J, Darzi A (2021) Reopening schools during the Covid-19 pandemic: governments must balance the uncertainty and risks of reopening schools against the clear harms associated with prolonged closure. Arch Dis Child 106(2):111–113
17. Szczygielski JJ, Bwanya PR, Charteris A, Brzeszczyński J (2021) The only certainty is uncertainty: an analysis of the impact of Covid-19 uncertainty on regional stock markets. Financ Res Lett 43:101945
18. Pawlak Z (2012) Rough sets: theoretical aspects of reasoning about data, vol 9. Springer Science and Business Media
19. Qin B, Zeng F, Yan K (2020) Uncertainty measurement for a tolerance knowledge base. Int J Uncertain Fuzziness Knowl-Based Syst 28(02):331–357
20. Ali G, Afzal M, Asif M, Shazad A (2021) Attribute reduction approaches under interval-valued q-rung orthopair fuzzy soft framework. Appl Intell, 1–26

21. Xue Y, Deng Y (2021) Decision making under measure-based granular uncertainty with intuitionistic fuzzy sets. Appl Intell, 1–10
22. Jain K, Kulkarni S (2020) Multi-reduct rough set classifier for computer-aided diagnosis in medical data. In: Advancement of machine intelligence in interactive medical image analysis. Springer, 167–183
23. Sowkuntla P, Sai Prasad PSVS (2021) Mapreduce based parallel fuzzy-rough attribute reduction using discernibility matrix. Appl Intell, 1–20
24. Sun B, Chen X, Zhang L, Ma W (2020) Three-way decision making approach to conflict analysis and resolution using probabilistic rough set over two universes. Inf Sci 507:809–822
25. Maldonado S, Peters G, Weber R (2020) Credit scoring using three-way decisions with probabilistic rough sets. Inf Sci 507:700–714
26. Bhapkar HR, Mahalle PN, Shinde GR, Mahmud M (2021) Rough sets in covid-19 to predict symptomatic cases. In: COVID-19: prediction, decision-making, and its impacts. Springer, pp 57–68
27. Liang J, Shi Z (2004) The information entropy, rough entropy and knowledge granulation in rough set theory. Int J Uncertain Fuzziness Knowl-Based Syst 12(01):37–46
28. Wei W, Liang J, Qian Y, Dang C (2013) Can fuzzy entropies be effective measures for evaluating the roughness of a rough set? Inf Sci 232:143–166
29. Düntsch I, Gediga G (1998) Uncertainty measures of rough set prediction. Artif Intell 106(1):109–137
30. Beaubouef T, Petry FE, Arora G (1998) Information-theoretic measures of uncertainty for rough sets and rough relational databases. Inf Sci 109(1–4):185–195
31. Wierman MJ (1999) Measuring uncertainty in rough set theory. Int J Gen Syst 28(4–5):283–297
32. Shah N, Ali MI, Shabir M, Ali A, Rehman N (2020) Uncertainty measure of z-soft covering rough models based on a knowledge granulation. J Intell Fuzzy Syst, (Preprint), 1–11
33. Li Z, Li Q, Zhang R, Xie N (2016) Knowledge structures in a knowledge base. Expert Syst 33(6):581–591
34. Wu W, Li H, Wang H, Zhu KQ (2012) Probase: a probabilistic taxonomy for text understanding. In: Proceedings of the 2012 ACM SIGMOD international conference on management of data, pp 481–492
35. Qian Y, Liang J, Dang C (2009) Knowledge structure, knowledge granulation and knowledge distance in a knowledge base. Int J Approx Reason 50(1):174–188
36. Li Z, Liu Y, Li Q, Qin B (2016) Relationships between knowledge bases and related results. Knowl Inf Syst 49(1):171–195
37. Qin B (2015) -Reductions in a knowledge base. Inf Sci 320:190–205
38. Sun W, Li J, Ge X, Lin Y (2021) Knowledge structures delineated by fuzzy skill maps. Fuzzy Sets Syst 407:50–66
39. Stefanutti L, Anselmi P, Chiusole DD, Spoto A (2020) On the polytomous generalization of knowledge space theory. J Math Psychol 94:102306
40. Shannon CE (2001) A mathematical theory of communication. ACM SIGMOBILE Mobile Comput Commun Rev 5(1):3–55
41. Yao Y (2003) Probabilistic approaches to rough sets. Expert syst 20(5):287–297
42. Qin B, Zeng F, Yan K (2018) Knowledge structures in a tolerance knowledge base and their uncertainty measures. Knowl-Based Syst 151:198–215
43. Kobren A, Monath N, McCallum A (2019) Integrating user feedback under identity uncertainty in knowledge base construction. Automated Knowl Base Const (AKBC)
44. Wu W, Zhang W, Li D, Liang J (2011) Theory and methods of rough sets. Chinese Scientific Publishers
45. Li J, Mei C, Lv Y (2011) Knowledge reduction in decision formal contexts. Knowl-Based Syst 24(5):709–715
46. Roy AG, Conjeti S, Navab N, Wachinger C (2019) Alzheimer's disease neuroimaging initiative et al. Bayesian quicknat: model uncertainty in deep whole-brain segmentation for structure-wise quality control. NeuroImage 195:11–22
47. Abdar M, Pourpanah F, Hussain S, Rezazadegan D, Liu L, Ghavamzadeh M, Fieguth P, Cao X, Khosravi A, Acharya UR et al (2021) A review of uncertainty quantification in deep learning: techniques, applications and challenges. Inf Fus 76:243–297