Cognitive Computation (2022) 14:1328–1349
https://doi.org/10.1007/s12559-022-09995-1

Concept Formation and Quantum-like Probability from Nonlocality in Cognition

Yukio-Pegio Gunji[1] **· Taichi Haruna**[2]

Received: 28 December 2020 / Accepted: 13 January 2022 / Published online: 20 April 2022

Abstract
Human decision-making is relevant for concept formation and cognitive illusions. Cognitive illusions can be explained by quantum probability, while the reason for introducing quantum mechanics is based on ad hoc bounded rationality (BR). Concept formation can be explained in a set-theoretic way, although such explanations have not been extended to cognitive illusions. We naturally expand the idea of BR to incomplete BR and introduce the key notion of nonlocality in cognition without any attempts on quantum theory. We define incomplete bounded rationality and nonlocality as a binary relation, construct a lattice from the relation by using a rough-set technique, and define probability in concept formation. By using probability defined in concept formation, we describe various cognitive illusions, such as the guppy effect, conjunction fallacy, order effect, and so on. It implies that cognitive illusions can be explained by changes in the probability space relevant to concept formation.

Keywords Decision-making · Conjunction fallacy · Quantum cognition · Concept formation · Lattice theory

Introduction

While human judgment [1] is individual, definite, and relevant to concept formation [2], decision-making tends to be stochastic [3]. Human judgment and decision-making have previously been studied independently. Formal concepts [4, 5] and rough sets [6, 7] have been used for concept formation [8–10]. Furthermore, models based on Bayesian inference [11–13] or quantum probability have been proposed [14–18] to explain cognitive illusions [19–24] derived from uncertainty in decision-making, and some have also explored the new research field of quantum cognition. While it is unclear that the reasons quantum mechanics can be applied to cognition have been verified [25], this topic has also been addressed under bounded rationality [26]. New challenges to concept formation and quantum probability are being discussed under the same framework [27, 28]; however, the relation between concept formation and quantum mechanics is still unclear.

In this study, we show that concept formation and quantum-like probability can be accounted for by the same framework, without the assumption of quantum mechanics based on axiomatic bounded rationality; the key idea is nonlocality in cognition. Human cognition is expressed as a relation between objects and attributes, and nonlocality in cognition is expressed as a specific relation. The relation is transformed into an ordered set of concepts (lattice) using a rough-set lattice technique. While the Boolean lattice corresponding to set-theoretic logic is obtained from an ideal relation, the orthomodular lattice corresponding to quantum logic is obtained from a relation with nonlocality. Probability is defined in concepts, and we show that a change in the probability space due to decision-making can describe various cognitive illusions. Although there have been some attempts to define probability in a lattice [29], our definition of probability explains cognitive illusions. Our approach can explore the generalization of nonlocality, which may involve quantum physics.

While we think that quantum cognition has the great ability to explain various cognitive illusions, we also think that quantum cognition is not necessarily needed to

✉ Yukio-Pegio Gunji
yukio@waseda.jp

1 Department of Intermedia Art and Science, School of Fundamental Science and Technology, Waseda University, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo 169-8555, Japan

2 Department of Information and Sciences, School of Arts and Sciences, Tokyo Woman's Christian University, 2-6-1 Zempukuji, Suginami-ku, Tokyo 167-8585, Japan

explain cognitive processes. The uncertainty of a thought is represented by superposition, multiple contexts are represented by Hilbert space, and the mixture of locality and non-locality is represented by entanglement in quantum theory [28]. However, we think that these features in the cognitive process can be represented without quantum theory. We here introduce the binary relation to represent cognitive process, obtain a non-Boolean lattice from the binary relation, and define the probability in a lattice. Since our model entails the uncertainty of a thought, multiple contexts and the mixture of locality and non-locality, it can explain cognitive illusions without quantum theory. The non-Boolean algebra we obtain is, however, connected to quantum information, because that is a quasi-disjoint system of Boolean algebras which is used to represent quantum logic in a lattice theory [30, 31].

Motivation of Incomplete Bounded Rationality

As mentioned before, quantum cognition of which quantum theory is used as mathematical tool or a kind of information theory has a great advantage to describe cognitive illusion. However, the question remains how the use of quantum theory in macroscopic phenomena such as cognitive process is verified. In this section, while we do not verify the use of quantum theory, we review and examine the previous attempt to verify applying quantum theory to cognitive process, reveal the weakness in that attempt, expand that attempt in overcoming the weakness, and construct new method to handle cognitive process. After that, we develop such a new method and define the framework by which cognitive illusion can be described.

There have been several attempts to show the potential reasons quantum mechanics can be applied to macroscopic phenomena such as cognition and/or decision-making [26, 31–35]. One attempt involves introducing Hilbert space without any assumption of quantum theory as physics [34]. Hilbert space is expressed by the direct sum of vector space and its orthogonal vector space. Because a vector in each space is expressed by a linear expression on an orthonormal basis, the space can be expressed as Boolean algebra whose atoms can correspond to the orthonormal basis. Boolean algebra implies a set lattice or a power set whose meet and join are defined by intersection and union, respectively. An element of the power set of an n-element set is expressed as an n-bit binary sequence, and each atom of the power set is expressed as a binary sequence in which only one digit is 1 and the other digits are 0. In Boolean algebra, an atom implies the occurrence of one element event in all element events. Thus, the probability of an atom is defined as $1/n$, and the sum of all atoms is $n*(1/n)=1$. In the case of a power set of a 3-element set, the atoms are 001, 010, and 100, and the probability is $P(001)=P(010)=P(100)=1/3$. Atoms can correspond to orthonormal bases entailing the additive measure function.

However, introducing Hilbert space does not directly imply quantum theory. Because the direct sum of orthogonal vector spaces just implies an almost disjoint union of Boolean lattices (i.e., the direct sum of some Boolean algebras except for the smallest and largest elements), there are no common elements between Boolean algebras except for the smallest and largest elements. For this reason, no element implies entanglement, which is characteristic of quantum theory.

Based on this point, previous studies have chosen the method to introduce additional mechanism to handle quantum theory, leading to a specific orthomodular lattice equipped with elements corresponding to entanglement. To clarify the weakness in that method, the simplest model [26, 35] is recalled here (Fig. 1).

Consider the situation in which a wandering firefly in box 1, 2, 3, or 4, as shown in Fig. 1, can blink. Assume that there are two directions of measurement, A and B. Since each measurement is restricted in terms of direction, an observer cannot indicate the accurate position of the firefly. In measurement A, measured location a reveals that the firefly is situated precisely in position, 1 or 3, and b reveals that it is situated precisely in position 2 or 4. The relation between the precise position and measured location is summarized in the above table in Fig. 1. The correspondence between the measured state and the location of the firefly is represented as $a(13)$ and $b(24)$. An assumption leading to an element with entanglement is the state in which the firefly does not blink. The "no blinking state" is indexed by 5 and is measured by n. In measurement A, the units of measurement are $a(13)$, $b(24)$, and $n(5)$. Thus, all statements containing measurements can be expressed as the power set of $\{a(13), b(24), n(5)\}$. This power set, which is 2^3-Boolean algebra, is shown in the Hasse diagram containing red elements in Fig. 1, where the empty set is represented by $\varnothing$. Similarly, in measurement B, the units of measurement are $c(12)$, $d(34)$, and $n(5)$. The power set with respect to measurement B is shown in the Hasse diagram containing blue elements in Fig. 1. It is easy to see that the two power sets have four common elements in terms of the precise location, such that $\varnothing$, $n(5)$, $ab(1234)=cd(1234)$, and $abn(12{,}345)=cdn(12{,}345)$. These common elements are represented by black elements. Unifying two power sets makes one orthomodular lattice containing blue, red, and black elements in Fig. 1. Except for the smallest and largest elements, the common elements reveal entanglement.

The question arises as to why "no blinking" is introduced. This might be the actual reason for implementing entanglement. The table below in Fig. 1 shows the relation between the precise position and measured location of a firefly not

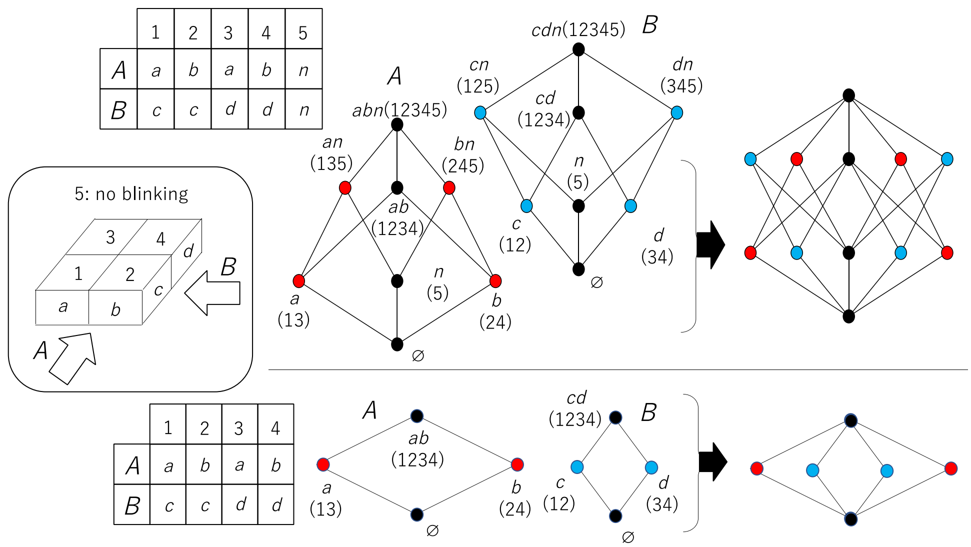

Fig. 1 Firefly model entailing two types of orthomodular lattices. The inner box represents the phenomenon in which a firefly blinks in any boxes 1–4 and the direction of measurement, *A* and *B*. The diagrams above represent constructing a lattice from the situation involving "no blinking." The diagrams below represent constructing a lattice (Chinese lantern) from the situation not involving "no blinking"

involving the no blinking state. In this condition, measurement *A* contains only two units of measurement, a(13) and b(24). Additionally, measurement *B* contains c(12) and d(34). Thus, each set consisting of statements containing measured states is expressed as 2^2-Boolean algebra. These two Boolean algebras have common elements, the smallest and largest elements. Combining two algebras with respect to common elements entails a Chinese lantern-type orthomodular lattice. Although this lattice also satisfies the definition of an orthomodular lattice and breaks the law of distributivity, there is no element corresponding to entanglement. Because it is just a Hilbert space, there is no advantage in constructing a firefly model if "no blinking" is not introduced. For this reason, no blinking is introduced, that is ad hoc assumption, and so is weakness in introducing an element representing entanglement.

We here focus on the weakness and examine the underlying framework of the firefly model. It results in finding the inevitable weakness in the framework. First, a fundamental question arises regarding the firefly model and why two measurements are prepared along orthogonal axes. This question is also applied to the foundation of Hilbert space and weak quantum logic [31–35]. They all introduce orthogonal axes, while their foundation is arbitrary. Although one of the hopeful motivations is bounded rationality [26], it introduces the sum of independent Boolean algebras. The rationality is expressed by the construction logic resulting from cognitive mapping. If cognitive map f is defined from universal set U to a set of representations, one can define an equivalence relation derived from this map and obtain a partition consisting of the equivalence classes with respect to the equivalence relation. Since each proposition is expressed as a combination of equivalence classes, the power set of a set of equivalence classes implies a whole set of propositions (Fig. 2A). This is simply Boolean algebra. Thus, the core of bounded rationality assumes that all the information in the world cannot be obtained and that it is therefore bounded; however, rationality is maintained in a restricted area. This implies that one can assign a cognitive map without ambiguity (Fig. 2A). How is this possible? It is assumed that the area outside of the bounded area is independently separated from the inside and that the inside is not influenced by the outside. This is a fundamental assumption. One measurement (i.e., observable) is chosen, the outside of the measurement is separated, and any state can be expressed as a union of atoms.

We think that fundamental assumptions cannot be established and is ad hoc. Even if one chooses an arbitrary observable, the bounded area cannot be separated from the outside. In other words, one cannot assign one map from a universal set to a set of representations. This is our basic assumption. Here, we show our roadmap to construct a quantum-like probability (Fig. 2B). We also assume bounded rationality, but it is incomplete. Incomplete bounded rationality is expressed as the ambiguity of a cognitive map, and then, it is expressed as two kinds of cognitive maps, f and g. From these two maps, one can obtain two kinds of equivalence relations and two kinds of partitions. Here, we introduce how to construct a lattice using the two partitions. A theory of expression in which any lattice can be obtained from two kinds of partitions has been verified [10]. Under this assumption, if $f = g$, a Boolean lattice is obtained; otherwise, a non-Boolean lattice can be obtained. The non-Boolean lattice contains not only a Chinese-lantern type orthomodular lattice but also an orthomodular lattice involving elements corresponding to entanglements. Thus, we introduce incomplete bounded rationality in our framework.

Here, we define the probability space for each lattice. This definition allows the probability space for 2^n-Boolean algebra, in which the probability of an atom is $1/n$, the

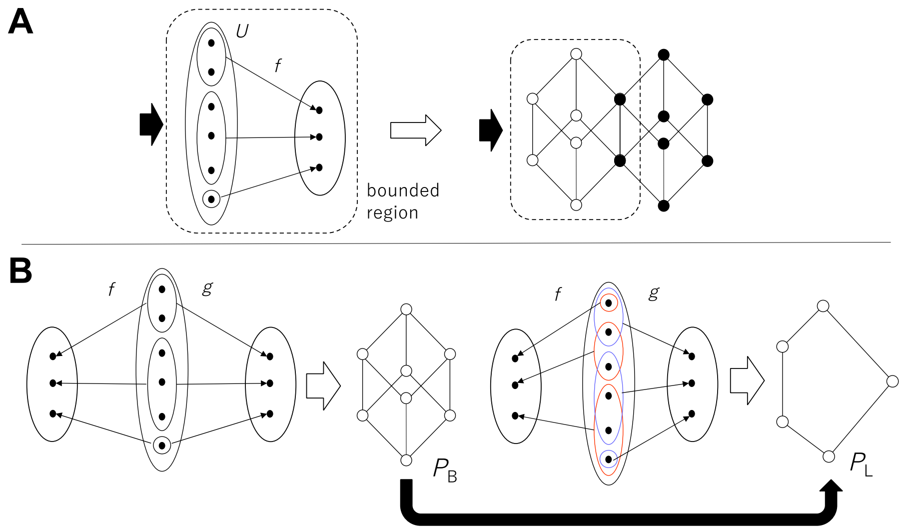

Fig. 2 **A** The bounded rationality in which cognitive map *f* can be assigned. This implies the almost disjoint union of Boolean algebras. **B** Incomplete bounded rationality. The ambiguity of the cognitive map is expressed as two kinds of maps. From them, two equivalence relations and two partitions are obtained, leading to both Boolean and non-Boolean lattices. In each lattice, one can define a probability space. Actual decision-making is expressed by changing the probability space

summation of the probability of all atoms is 1, and the probability satisfies the additive function. The probability space of Boolean algebra is represented by P_B. The probability derived from a non-Boolean lattice is represented by P_L. Thus, our probability model allows the extension of quantum probability on an orthonormal basis. Since Boolean algebra corresponds to the terminal element (i.e., complete bounded rationality) in the spectrum of incomplete bounded rationality under our framework, Boolean algebra implies an ideal case that is different from actual cases. In actual cases, there are ambiguities in the cognitive map, and a non-Boolean lattice is obtained. We think that actual decision-making implements the combination of P_B and P_L. Before facing an actual situation, one can calculate the probability of an event in P_B. That is ideal situation. Under the actual situation, the cognitive map has ambiguity, one is under a non-Boolean lattice, and one calculates the probability of an event in P_L. While quantum cognition compares ideal inference to the inference in classical probability and empirical inference to the inference in quantum probability. Similarly, our model without quantum theory compares ideal inference to the inference in P_B and empirical inference to the inference in P_L. Because our actual inference is implemented in the empirical world consisting of multiple contexts, such a comparison is reasonable.

Under incomplete bounded rationality, one can obtain an orthomodular lattice. There is, however, a major difference with respect to constructing a sum of Boolean lattices under complete bounded rationality and our framework under incomplete bounded rationality. In complete bounded rationality, the outside of each bounded region cannot influence the region itself. A whole universe consisting of multiple bounded regions can result in an orthomodular lattice (Fig. 2A). By contrast, in our framework, even in a single bounded region, an orthomodular lattice can be obtained due to the influence of the outside of the bounded region. In our framework, the ambiguity of a cognitive map is expressed as correlation between the inside and outside of a context, that is called "nonlocality" in cognition.

There have been other attempts to construct orthomodular lattice-containing elements corresponding to entanglements. Such attempts are approaches using category theory. One adds the tensor product as a bilinear operation with a category of Hilbert space [36]. Another introduces nonlocality [37]. While they are axiomatic, our approach is intuitive because the implication of "nonlocality" is different from quantum mechanical reasoning.

Cognitive Illusions and Concept Formation

Although humans have previously been considered to be rational decision-makers, many cognitive and linguistic experiments have shown that humans make irrational decisions, which is inconsistent with probability theory [1, 13, 38–40]. Such decisions are called cognitive illusions. Since these decisions are made with respect to an event or concept, they are relevant for concept generation. Thus, cognitive illusions and concept generation should be studied together. Our approach is the first attempt to provide an explanation for cognitive illusions in the framework of concept formation. To indicate the significance of our approach, we here discuss cognitive illusions and concept formation.

Cognitive Illusions

Cognitive illusions have been experimentally verified by evaluating the frequency of decision-making [41]. The joint probability of multiple events is equal to or less than that of a single event. Against probability theory, reversed inequalities are found in human cognition, such as in the conjunction fallacy and guppy effect [15, 16, 19, 24, 42, 43]. The conjunction fallacy implies

$$P(A \text{ AND } B) > P(A). \tag{1}$$

It is clear that this inequality is the reverse of that of joint probability in probability theory. The guppy effect is a special case of the conjunction fallacy [16, 24]. The probability of thinking about a guppy with respect to "fish" is represented by P_{guppy} (fish), and the probability of thinking about a guppy with respect to "pet" is represented by P_{guppy} (pet). Because people usually recall tuna or salmon when asked to provide an example of fish, P_{guppy} (fish) is very small. Similarly, P_{guppy}(Pet) is very small. In contrast, a guppy is frequently recalled when one is asked for an example of a pet fish. This leads to

$$P_{\text{guppy}}(Pet \text{ AND } Fish) \geq (P_{\text{guppy}}(Pet) + P_{\text{guppy}}(Fish))/2. \tag{2}$$

This implies that there are some A and B such that $P(A \text{ AND } B) \geq (P(A) + P(B))/2$. This is the general form of the guppy effect [19, 24, 42].

The order effect is also a cognitive phenomenon that is inconsistent with probability theory [44, 45]. In probability theory, $P(A \text{ AND } B) = P(B \text{ AND } A)$. Imagine that you love hot Indian dishes. The statement "A AND B" is frequently expressed by "A but B" in everyday life. For you, Indian dishes are "hot but good" rather than "good but hot." Thus, you are more likely to describe the dish as "hot but good" than as "good but hot." This situation can be generalized by

$$P(A \text{ AND } B) \neq P(B \text{ AND } A). \tag{3}$$

This implies that the first term modifies the significance of the second term.

The Ellsberg and Machina paradoxes are also inconsistent with probability theory, which implies

$$P(A) = P(A \text{ AND } B) + P(A \text{ AND NOT } (B)). \tag{4}$$

Humans sometimes violate this equation in decision-making. Ellsberg [46] proposed the following thought experiment, and Aerts and others [47] conducted cognitive experiments with over one hundred participants (also see [48]). Imagine 90 colored balls are contained in an urn that contains 30 red balls, and the sum of yellow and black balls is 60, where the proportion of yellow and black balls is unknown. One ball is randomly drawn from the urn. If the participant predicts the color of the drawn ball, he/she obtains a reward. The results of the experiment show that most participants bet on red balls.

Although the probability of a black ball being pulled is equal to that of a red ball being pulled, participants preferentially predict that a red ball will be drawn. How can we explain this preference? Let the event of drawing black balls and that of drawing yellow balls be A and B, respectively. Although $P(A) = 1/3$, uncertainty regarding the proportion of black and yellow balls could contribute to a cognitive fallacy in calculating the probability. Imagine that two balls are drawn from a set of 60 black and yellow balls, where one ball is black. Thus, there are two possibilities such that two balls are "black (A) and yellow (B)" or "black (A) and nonyellow (NOT(B))," which is expressed as

$$P(A) \neq P(A \text{ AND } B) + P(A \text{ AND NOT } (B)). \tag{5}$$

This inequality underlies the Ellsberg and Machina paradoxes.

Borderline contradiction implies that the statement "Tom is fat and not fat" is sometimes true, dependent on contexts [21, 49, 50]. If Tom is in a training gym, he is fat. By contrast, if he is in a room with sumo wrestlers, he is not fat. Thus, it is generalized by.

$$P(A \text{ AND NOT } (A)) > 0. \tag{6}$$

Borderline contradiction is relevant for essential vagueness in cognition [51]. It has been concluded that vagueness can result from super- and sub-interpretation. This idea is directly expressed by the rough-set lattice mentioned later. Although vagueness had been expressed using membership functions in fuzzy logic, it was recently replaced by a rough set of lower and upper approximations [6–10].

Cognitive illusions have recently been explained using quantum mechanics in the research field called quantum cognition [14, 16–18, 23, 24, 47, 48, 52, 53]. Since quantum probability allows joint probability to entail interference terms, various cognitive illusions can be explained using such terms. However, there is no physical reason to apply quantum mechanics to cognitive processes [25]. As mentioned in "Motivation of incomplete bounded rationality," there are some attempts to verify the foundation to apply quantum theory to cognitive process; it remains unclear. Our examination leads to incomplete bounded rationality, which is defined later.

Concept Formation

As mentioned before, cognitive illusions are estimated by the probability of events generated through cognitive processes in everyday life. Events have previously been studied as concepts and/or semantically quantized things in the field of information science, while events have been studied as categories in the field of cognitive linguistics. These studies are sometimes inconsistent with each other. Here, we develop a definite concept that is consistent with both concepts in information science and categories in cognitive linguistics.

Concepts in information science can be defined in terms of extent and intent [4]. Given a world consisting of a set of objects, G, attributes, M, and the relation between G and M, $I \subseteq G \times M$ is a formal concept generated by collecting objects and attributes. If object $g \in G$ has a relation to attribute $m \in G$, it is expressed as gIm. For a set of objects $A \subseteq G$, a set of collected attributes is defined by

$$A' = \{m \in M | gIm, {}^{\forall} g \in A\} \tag{7}$$

Similarly, for a set of attributes $B \subseteq M$, a set of collected objects is defined by

$$B' = \{g \in G | gIm, {}^{\forall} m \in B\} \tag{8}$$

Under this definition, the pair (A, B) such that $A' = B$ and $B' = A$ is a formal concept, and A and B are the extent and intent, respectively. It is easy to see that each element of extent satisfies all elements of intent and that each element of intent can be applied to all elements of extent.

It is clear to see that collecting objects and attributes in (7) and (8) implies an approximated set. In this sense, one can compare the formal concept to the concept approximated as a rough set. The notion of a rough set is defined by the following [6–10]. Given universal set U, if there exists equivalence relation R in U (i.e., $R \subseteq U \times U$), the universal set is divided into a disjoint partition called the equivalence class of R. For element y in U, the equivalence class is defined as $[y]_R = \{x \in U | xRy\}$, where xRy implies that x is equal to y with respect to R. Using an equivalence class, one can define two kinds of rough sets for a subset X of U as $R_*(X)$ and $R^*(X)$, which are given by

$$R_*(X) = \{x \in U | [x]_R \subseteq X\}, \tag{9}$$

$$R^*(X) = \{x \in U | [x]_R \cap X \neq \emptyset\}. \tag{10}$$

It can be easily seen that $R_*(X) \subseteq X \subseteq R^*(X)$, which implies that $R_*(X)$ and $R^*(X)$ are necessary and sufficient conditions for X. Thus, set X satisfying $R_*(R^*(X)) = X$ implies a set satisfying the necessary and sufficient condition for X.

There are two ways to collect elements, intent (attributes) and extent (objects), in formal concepts. Here, we introduce two ways to collect elements, objects, and attributes, in rough sets using two kinds of equivalence relations, R and S, respectively. Each object is defined by an equivalence class of R such as $[x]_R$, and each attribute is defined by an equivalence class of S such as $[x]_S$. Thus, one can have two kinds of rough sets with respect to R and S: R_*, R^*, K_*, and K^*. Therefore, the necessary and sufficient condition is formed by combining two kinds of relations, such as

$$S_*(R^*(X)) = X. \tag{11}$$

A solution (or a fixed point) satisfying Eq. (11) is a set of attributes. Figure 3 shows the significance of using two kinds of equivalence relations. For a universal set, two kinds of partitions are derived from equivalence relations S and R (Fig. 3A). Each part of a universal set represents the

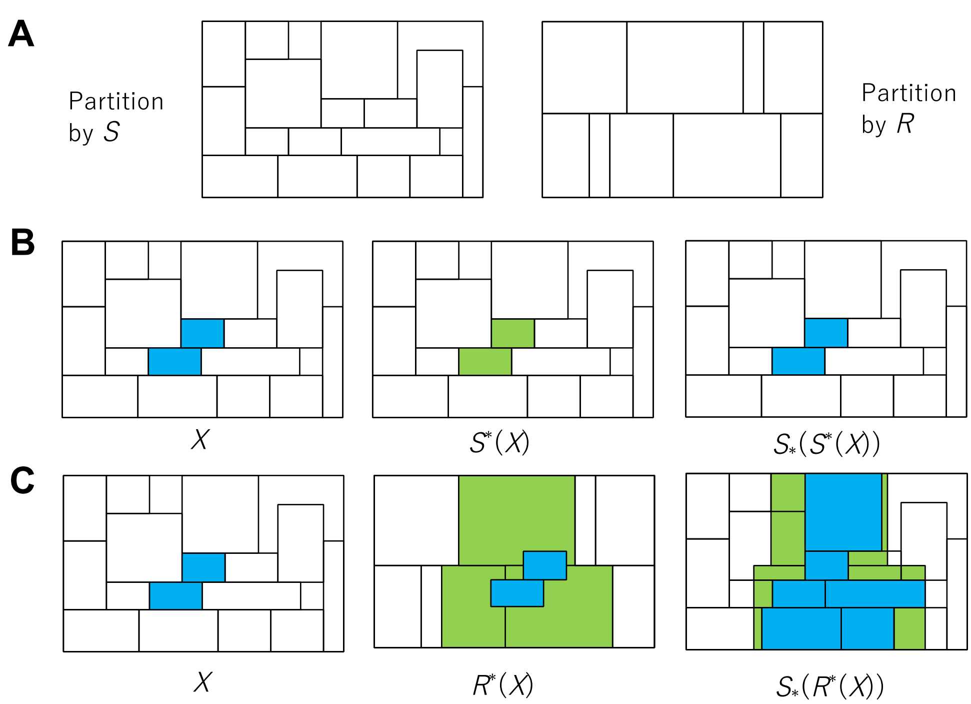

Fig. 3 The significance of two kinds of partitions for the necessary and sufficient condition of rough set. **A** Two kinds of partitions are defined by equivalence relation S and R for a universal set. Each part represents an equivalence class. **B** If only S is used for the necessary and sufficient condition, any union of equivalence classes satisfies $S_*(S^*(X)) = X$. **C** If two kinds of equivalence relations are introduced, some unions of equivalence classes do not satisfy $S_*(R^*(X)) = X$

equivalence class. Figure 3B shows $S_*(S^*(X)) = X$. If only one equivalence relation is introduced (e.g., S) to define Eq. (11), any union of the equivalence classes, X, can be a solution to $S_*(S^*(X)) = X$. Here, X is painted blue, $S^*(X)$ is painted green, and $S_*(S^*(X))$ is painted blue in Fig. 3B. By contrast, if two kinds of partitions by S and R are introduced as necessary and sufficient conditions, there are some X that do not satisfy $S_*(R^*(X)) = X$. Figure 3C shows such a case, where X, which is a union of equivalence classes of S, is painted blue (left), $R^*(X)$ is painted green (middle and right), and $S_*(R^*(X))$ is painted blue (right).

Compared to the formal concept, X satisfying condition (11) implies a concept derived from a rough set. The next question is whether a formal concept or a concept derived from a rough set is more likely to be the linguistic concept used in everyday life.

In cognitive linguistics, the concepts that people generate and use in everyday life are called categories. While a category is also constructed from a pair of objects and attributes, it is different from a formal concept with respect to the extent of heterogeneity [19, 20, 40, 54, 55]. The objects in a category do not have common attributes. In the category of birds, although representative objects called prototypes, such as sparrows, have the representative attribute of flying, penguins do not have this attribute. Such a representative object is called a prototype. It is reported that there is family resemblance among the objects in a category. For instance, a father has a nose similar with that of his elder son but not that of his daughter. In contrast, a mother's ears are similar with those of her daughter. In this manner, the objects in a category only partially share common properties. That is, family resemblances constitute categories.

Due to the nature of prototypes and family resemblance, a category or a set of objects has a heterogeneous structure. In this sense, a category or concept observed in cognitive linguistics is totally different from a formal concept. By contrast, a concept derived from a rough-set has heterogeneity. If the equivalence relation of objects is used as the necessary condition, the equivalence relation of attributes is used as the sufficient condition, which results in a heterogeneous structure in which some objects have more attributes and other objects have few attributes. Such heterogeneous structures are always found in concepts derived from rough sets. For this reason, we use concepts derived from a rough set instead of formal concepts.

We here repeat that we do not apply quantum theory to cognitive process. Most of Eqs. (1)–(6) are just fact in cognitive experiments, outside the classical probability calculation. Although quantum cognition applied quantum theory to cognitive process to describe cognitive illusions, we here attempt to describe cognitive illusion without cognitive theory. Our model is discussed in the next section.

Methods

Nonlocality in Cognition

Nonlocality in the macroscopic world has previously been discussed with respect to the violation of Bell inequality [56]. Our idea is consistent with previous ideas and is generalized in terms of the relation between objects and attributes in cognitive processes. First, our idea is defined, and finally, we argue the relationship between our "nonlocality" and Bell inequality in the macroworld.

As mentioned before, decision-making is performed using concepts defined by the binary relations between objects and attributes. In the tradition of quantum logic or orthomodular lattice representing quantum logic, entanglement corresponding to the nonlocality (or the mixture of locality and nonlocality) is expressed as a specific element in orthomodular lattice. In that representation, an element featured with nonlocality is expressed as an element belonging to multiple contexts (i.e., multiple Boolean algebras). It implies that nonlocality can be expressed as combination of "belonging to" and "not belonging to" a context, and that nonlocality can be expressed by binary relation. Recall the category bird, whose main attribute is flying. As concrete objects, sparrows are judged to be birds or not by estimating whether sparrows satisfy the attribute of flying. Thus, nonlocality in cognition might be defined by using the relation between objects and attributes.

Because nonlocality plays an essential role in cognitive illusions, quantum mechanics has been introduced to explain it. It has been debated whether nonlocality is possible in macroscopic cognitive processes. Nonlocality in quantum physics implies that "measuring the state of an object influences the state of other objects that are located far from the measured object." In this study, we propose nonlocality in cognition, where nonlocality is extended in terms of the notion of space. Although nonlocality in cognition looks different from nonlocality in quantum theory, inevitable coexistence of A and non-A (nonlocality in cognition) cannot be distinguished from nonlocality of which state A at the one site leads to non-A at the other site. In extending nonlocality, distance, implying "far from the measured object," is estimated not in physical space but in concept space or neural information space [28, 60–62]. Here, we define nonlocality by global correlation (relation) in concept space expressed by a binary relation.

We take the example of the cognition of a cat, as shown in Fig. 4. Herein, we use words that start with uppercase letters for the names of objects, i.e., cat, dog, and so on, and words starting with lowercase letters to denote attributes. One can determine whether the object (Cat) is a cat or not. This object may appear to be a cat; however, it may

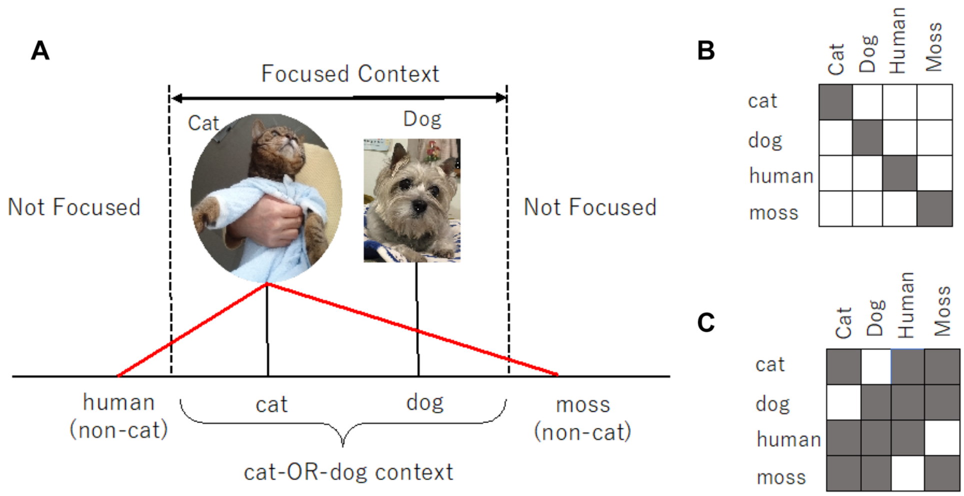

Fig. 4 **A** "Cat rather than noncat" expressed in a focused context. Determining a given object, cat, as a cat contains not only the attributes of a cat but also the attributes of a noncat. The ambiguity of cat and noncat can help determine a cat only by indicating a specific context, the cat-OR-dog context. **B** A relation featuring "cat" without context. **C** A relation featuring "cat rather than noncat" using the cat-OR-dog context

also appear to be a human, a mass of dust, and/or moss. Thus, cat is sequentially determined by.

cat (because it has a typical tabby pattern).
noncat (because it wears clothes like a human).
cat (because it has a long tail).
noncat (because its fur is like moss).

This observation implies that the object is both a cat and a noncat, and this directly implies undecidability. However, one can determine whether an object is either a cat or a noncat in everyday life. For example, this object can be determined to be a cat because the attributes of a cat realized by the object are considerably more important than those of a noncat realized by this object. This thinking process is reasonable. However, the importance of an attribute depends on a finite set of attributes measured by the observer (i.e., the sequence of cat and noncat mentioned above is finite). If one considers the "next" attribute to the attribute, the "like moss" nature of an object, one may find essential noncat attributes that are more important than the cat attributes, and the object can be determined to be a noncat. Thus, although one can avoid undecidability, the decision is intrinsically arbitrary. Listing the attributes of a cat or a noncat is performed as if the final decision were made before listing the attributes.

While this object can be judged to be both a cat and a noncat, the observer decides that it is a cat rather than a noncat. The statement "cat rather than noncat" is strongly relevant for the law of reasoning. The reasoning involved in deciding that the object is a cat is nearly equal to that for deciding it is a noncat. The question arises, whether coexistence of cat and noncat is accidental or not. We will answer this question.

How can one guarantee the condition of "cat rather than noncat"? To do so, we introduce the context that one focuses on during decision-making. Figure 4A shows that an object can be identified as a cat by distinguishing it from a dog. In the focused cat-OR-dog context, a cat has no noncat attributes, i.e., attributes of dogs. However, a cat has noncat attributes that are outside of the context of focus. The relationships between objects and attributes can be expressed as binary relations in which the relationship between an object and an attribute either exists or not.

When there are two contexts, such as cat-OR-dog and human-OR-nature, as shown in Fig. 4C, each context can be expressed as a diagonal matrix in which the attribute and object are related only by diagonal elements, and the attributes and objects may have relationships outside these contexts. In this matrix, each cell represented by an (object, attribute) pair is shaded if the object is related to the attribute, and a pair is blank if the object is not related to the attribute.

Introducing a context, one can define nonlocality in cognition. In a specific context, cat may be distinguished from other objects, whereas outside the context, cat may have overlapping characteristics with other objects. It may be related to the attribute of a cat rather than the attribute of a noncat (Fig. 4C). While cat has no relation to noncat in the cat-OR-dog context, cat has a relation to all attributes (i.e., noncat attributes) outside the cat-OR-dog context. This implies that "when the cat is determined as a cat, this decision is correlated to any other attributes far from the cat-OR-dog context." This is nonlocality in cognition, where "far from" is defined not in physical space but in concept or information space. While decisions such as "cat rather than noncat" can be made with such relationships, there is a possibility of misunderstanding that "cat" is identified with "noncat." Therefore, we can conclude that coexistence of cat and noncat is not accidental but inevitable. If an object is determined as cat in the context, then the attribute outside the context is destined to be noncat. Thus, the situation is not distinguished from the situation in which the state at some place influences the state at any other place. That is why we call it nonlocality in cognition. By contrast, a relation without any ambiguity can be expressed as a diagonal matrix, as shown in Fig. 4B.

We consider relationships without ambiguity and those with multiple contexts as being ideal and actual, respectively. Concrete and real decision-making is always achieved by introducing a focused context, which can then be expressed as a relationship with multiple contexts (Fig. 4C); meanwhile, single objects are expressed in relationships without ambiguity in the ideal case because they are identified by their own attributes (Fig. 4B).

This idea entails that the relationships between objects and attributes can change before and after actual logical operations. Before applying a logical operation, one uses the matrix without context, and after applying a logical operation, the matrix may be used with context. Owing to this shift, the logic and the probability space can be changed to encompass cognitive illusions. We thus define the logical space and probability space derived from a matrix or a binary relation, as mentioned in the next section.

Finally, we discuss the relationship between previous arguments on nonlocality in the macroworld and our nonlocality. Violation of Bell inequality is generalized in the macroworld in terms of combinations of four experiments, e_1, e_2, e_3, and e_4 [51]. When a pair of experiments e_i and e_j can be performed together, this is represented by e_{ij}. The performance of each experiment is estimated by two values, up and down, and the coincidence of two experiments, e_i and e_j, is estimated by expected value $\mathbf{E}_{ij}$, which is represented by $+1$ corresponding to the coincidence of two experiments, (e_i = up and e_j = up) or (e_i = down and e_j = down), and -1 corresponding to the noncoincidence of two experiments, (e_i = up and e_j = down) or (e_i = down and e_j = up). If the macroscopic experiments are set up under this framework, Bell inequality is defined by [51]

$$|\boldsymbol{E}_{13} - \boldsymbol{E}_{14}| + |\boldsymbol{E}_{23} + \boldsymbol{E}_{24}| \leq 2. \tag{12}$$

There are some experimental setups that can violate Bell inequality, which implies nonlocality in the macroworld [51]. It is easy to see that our definition of nonlocality also violates Bell inequality. Recall the binary relation between objects and attributes shown in Fig. 1C. Here, we define four experiments by assigning attributes, cat (e_1), dog (e_3), human (e_2), and moss (e_4), and the performance of the experiment is estimated by the relation between attributes and objects, cat and dog (Fig. 4C) such that

$$\begin{array}{lcc} & \text{Cat} & \text{Dog} \\ \text{cat } (e_1) & 1 & 0 \\ \text{dog } (e_3) & 0 & 1 \\ \text{human } (e_2) & 1 & 1 \\ \text{moss } (e4) & 1 & 1. \end{array} \tag{13}$$

The relation between an attribute and an object is defined to assign the connection of the performance, $e_i = 1$ implies the connection opened to other experiments, and $e_i = 0$ implies no connection. If two experiments are opened to each other, they can coincide with each other. Thus, two experiments, e_i and e_j, can coincide with each other if $e_i = e_j = 1$ in some objects. The coincidence of a pair of experiments is estimated by "cat or dog." Two experiments performed together are verified to coincide with each other, and the expected value of a pair of experiments is assigned by $+1$ if $e_i = e_j = 1$ in cat or dog. Similarly, two experiments are verified not to coincide with each other, and the expected value is assigned to be -1 if the condition, $e_i = e_j = 1$ never holds in cat or dog. For example, e_1 and e_3 are different with respect to both cat and dog; thus, $\mathbf{E}_{13} = -1$. By contrast, e_2 and e_3 have the value of 1 with respect to dog, and then, they satisfy the condition that they coincide with each other with respect to cat or dog. Thus, $\mathbf{E}_{23} = +1$. Similarly, $\mathbf{E}_{14} = +1$ because $e_1 = e_4 = 1$ in cat, and $\mathbf{E}_{24} = +1$ because $e_2 = e_4 = 1$ in both cat and dog. This leads to

$$|\boldsymbol{E}_{13} - \boldsymbol{E}_{14}| + |\boldsymbol{E}_{23} + \boldsymbol{E}_{24}| = |-1-1| + |1+1| = 4. \tag{14}$$

This implies that the four experiments violate Bell inequality.

Under our approach, nonlocality essentially influences violation of Bell inequality. We define nonlocality as an overflow of the relation (information) beyond the context. Thus, if the nonlocality is removed from the relation, $e_2 = e_4 = 0$ in both cat and dog. Thus, $\mathbf{E}_{13} = \mathbf{E}_{23} = \mathbf{E}_{14} = \mathbf{E}_{13} = -1$, and then, $|\mathbf{E}_{13} - \mathbf{E}_{14}| + |\mathbf{E}_{23} + \mathbf{E}_{24}| = |-1+1| + |-1-1| = 2 \leq 2$. Thus, Bell inequality is not violated. In our framework of nonlocality, decision-making is performed if the ambiguity of the decision is hidden outside the context, which implies that the decision on a specific concept (i.e., an object-attribute pair) is globally related to any other concept. This idea is consistent with global work space theory supported in brain research [57–59].

Lattice Structure Corresponding to a Binary Relation

Given a binary relation, one can construct a logical structure called a lattice using a rough-set lattice technique. Lattice L is a partially ordered set that is closed with respect to two logical operations: AND and OR. The term "closed" implies that, for any two elements, X and Y of a lattice, X AND Y and X OR Y are also elements of the same lattice. The definitions of the operations and representations of a lattice are described in detail in Appendix 1.

Using a rough set, we define a set of concepts using Eq. (11). In our framework, there are two equivalence relations, R and S (see Appendix 2). Because it is impossible to choose either R or S as a necessary and sufficient condition, we use one as a sufficient condition and the other as a necessary condition. It has previously been verified that a collection of X such that $R_*(S^*(X))=X$ is a lattice [8–10]. Set X satisfying this equation is a concept in a rough-set lattice. It has also been verified that lattices obtained from $R_*(S^*(X))=X$, $S_*(R^*(X))=X$, $R^*(S_*(X))=X$ and $S^*(R_*(X))=X$ are isomorphic to each other [10]. Here, we construct a lattice as follows:

$$L = \{X \subseteq U | S_*(R^*(X)) = X\}. \quad (15)$$

As mentioned before, the independent items are taken in a specific binary relation, a diagonal matrix, as shown in Fig. 4B. After the logical operations in actual situation, the binary relation is expressed as diagonal matrices with multiple contexts, as shown in Fig. 4C. Next, we address the type of lattice that can be obtained. We show that a relation in the form of a diagonal matrix without multiple contexts entails a Boolean lattice and that a relation with multiple contexts entails an orthomodular lattice.

Figure 5A shows 2 by 2 to 6 by 6 diagonal matrices and corresponding lattices expressed as Hasse diagrams (see Appendix 1). In an n by n diagonal matrix, the objects are expressed as $A_1, A_2, \ldots, A_n$, and the attributes are expressed as $a_1, a_2, \ldots, a_n$. It is assumed that the objects and attributes are equivalence classes with respect to equivalence relations R and S, respectively. It can be clearly seen that, for any subset of a universal set consisting of attributes, $S_*(R^*(a_k \cup a_p \cup \ldots \cup a_q)) = S_*(A_k \cup A_p \cup \ldots \cup A_q) = a_k \cup a_p \cup \ldots \cup a_q$. Thus, any union of attributes is an element of a lattice, $L=\{X \subseteq U | S_*(R^*(X))=X\}$. The lattice contains all attributes as atoms (i.e., with the nearest element being larger than the smallest element) and all unions of atoms. This implies that the lattice is a power set of attributes and that it is a Boolean lattice.

Figure 5B shows a typical binary relation consisting of some subdiagonal matrices. A given n by n relation is divided into parts. Without loss of generality, we discuss how an orthomodular lattice is obtained. A 10 by 10 relation consists of one 4 by 4 and two 3 by 3 diagonal matrices, as shown in Fig. 5B. Let $a_1, a_2, \ldots, a_{10}$ and $A_1, A_2, \ldots, A_{10}$ represent attributes and objects, respectively. The 4 by 4 matrix consists of $(A_i, a_j)=1$ if $i=j$ and $(A_i, a_j)=0$ otherwise, where the value 1 implies that an attribute is related to an object, 0 represents no relation, and $i, j=1, 2, \ldots, 4$. For $a_k \cup a_p \cup a_q$ with $k, p, q=1, 2, \ldots, 4$, it can be clearly seen that $S_*(R^*(a_k \cup a_p \cup a_q)) = S_*(A_k \cup A_p \cup A_q \cup (A_5 \cup A_6 \cup \ldots \cup A_{10})) = a_k \cup a_p \cup a_q$ and that $S_*(R^*(a_1 \cup a_2 \cup a_3 \cup a_4)) = S_*(U) = U$. This implies that all unions of attributes in $\{a_1, a_2, a_3, a_4\}$ are elements of the rough-set lattice, L, except $a_1 \cup a_2 \cup a_3 \cup a_4$.

This implies that the 2^4-Boolean lattice, except for the largest element, is a sublattice in L. Every union of attributes that belongs to different diagonal matrices, such as $a_3 \cup a_6$, is not an element of L because $S_*(R^*(a_3 \cup a_6)) = S_*(U) = U$. This implies that one 2^4-Boolean lattice and two 2^3-Boolean lattices can constitute a disjoint union except for the smallest and largest elements (i.e., an almost disjoint union of Boolean algebras) [30, 60]. The Hasse diagram of this almost disjoint union of sub-Boolean lattices is shown in Fig. 5B. It can be directly verified that most of the almost

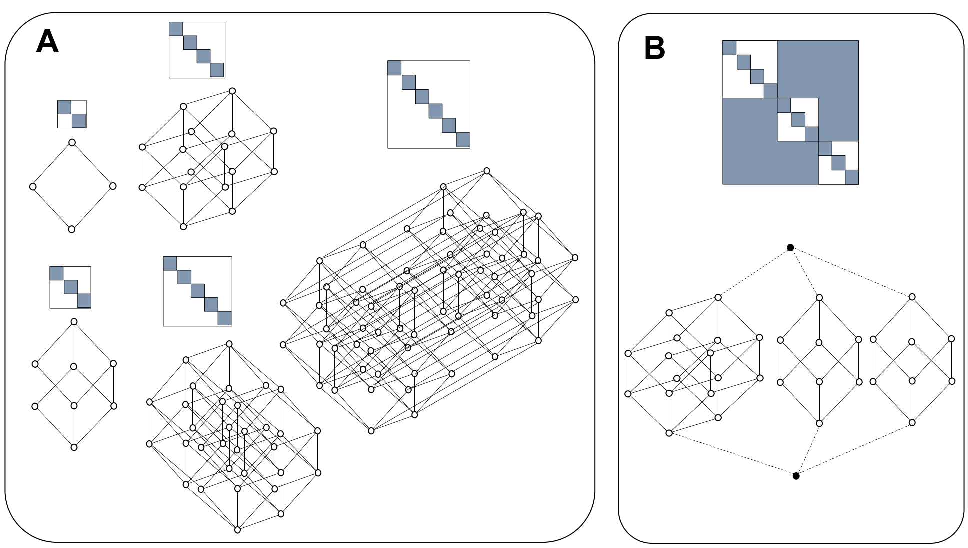

Fig. 5 **A** Boolean lattices obtained from diagonal matrices. If a diagonal matrix is an n by n matrix, a 2^n-Boolean lattice (i.e., a lattice consisting of 2^n elements) is obtained. Lattices are drawn as Hasse diagrams. **B** An orthomodular lattice is obtained from diagonal matrices with multiple contexts. Each n by n sub-diagonal matrix leads to a 2^n-Boolean lattice, and an entire lattice is expressed as an almost disjoint union of some Boolean sub-lattices. A lattice is expressed as a Hasse diagram except for the smallest and largest elements, where the smallest and largest elements (black circles) are common to all Boolean sub-lattices

disjoint unions of Boolean algebras are an orthomodular lattice [60, 63–65]. Finally, the logical structure before the logical operation is expressed as a Boolean lattice, and the logical structure after the operation is expressed as an orthomodular lattice. Figure 6 shows the flow chart of constructing a lattice for a given binary relation.

Probability in Lattice

Here, the N-bit Boolean lattice is represented by $\mathbf{B}_N$, which consists of binary sequence $A = a_1a_2 \ldots a_N$, where a_k is either 0 or 1. We consider L_N, a subset of $\mathbf{B}_N$, where L_N is a lattice with the smallest and largest elements of $\mathbf{B}_N$. An element of L_N, A, is ordered by natural number i and is represented by A_i such that

$$j < i \quad \text{if } \mathbf{1}(A_i) < \mathbf{1}(A_j); \tag{16}$$

$$j \le i \quad \text{if } \mathbf{1}(A_i) = \mathbf{1}(A_j) \text{ and } \Omega(A_i) \le \Omega(A_j). \tag{17}$$

where $\mathbf{1}(A)$ is the number of 1 s in binary sequence A and $\Omega(A)$ represents the decimal number corresponding to A. For example, $\mathbf{1}(001001) = 2$ and $\Omega(001001) = 9$. Given $L_3 = \{111, 110, 010, 001, 100, 000\}$, the elements of L_3 are ordered by $A_1 = 111$, $A_2 = 110$, $A_3 = 100$, $A_4 = 010$, $A_5 = 001$ and $A_6 = 000$. Given L_N, $A = a_1a_2 \ldots a_N$ and $B = b_1b_2 \ldots b_N$ in L_N are binary sequences, and order $A \le B$ is defined by $a_k \le b_k$ for any k. The AND and OR operations are defined by

$$A \text{ OR } B = \min\{X \in L_N | A \le X, B \le X\} \tag{18}$$

$$A \text{ AND } B = \max\{X \in L_N | X \le A, X \le B\}. \tag{19}$$

Under this definition, if A AND B and A OR B are also in L_N for any A, B in L_N, then L_N is a lattice. We also define NOT(A) as a binary sequence satisfying

$$A \text{ OR NOT } (A) = \mathbf{1} \quad \text{and} \quad A \text{ AND NOT}(A) = \mathbf{0}, \tag{20}$$

where $\mathbf{0}$ and $\mathbf{1}$ represent the smallest and largest elements in a lattice, respectively.

For A_i in L_N, the probability of A_i, represented by $P(A_i)$, is defined by $Prob([A_i]_L)$, where $Prob$ is defined later and $[A_i]_L$ is a subset (i.e., partition) of $\mathbf{B}_N$ defined as follows:

$$[A_i]_L = \{A_j \in \mathbf{B}_N | A_i \le A_j\} - \cup_{j<1} [A_j]_L \tag{21a}$$

Because $\mathbf{B}_N$ is divided into disjoint subsets with respect to L_N, for any A in $\mathbf{B}_N$,

$$[A]_L = [A_i]_L, \quad \text{if } A \in [A_i]_L. \tag{21b}$$

Consider $L_3 = \{111, 110, 100, 010, 001, 000\}$, we have $[A_1]_L = \{111\}$, $[A_2]_L = \{110\}$, $[A_3]_L = \{100, 101\}$, $[A_4]_L = \{010, 011\}$, $[A_5]_L = \{001\}$, and $[A_6]_L = \{000\}$. In other words, $\mathbf{B}_N$ is

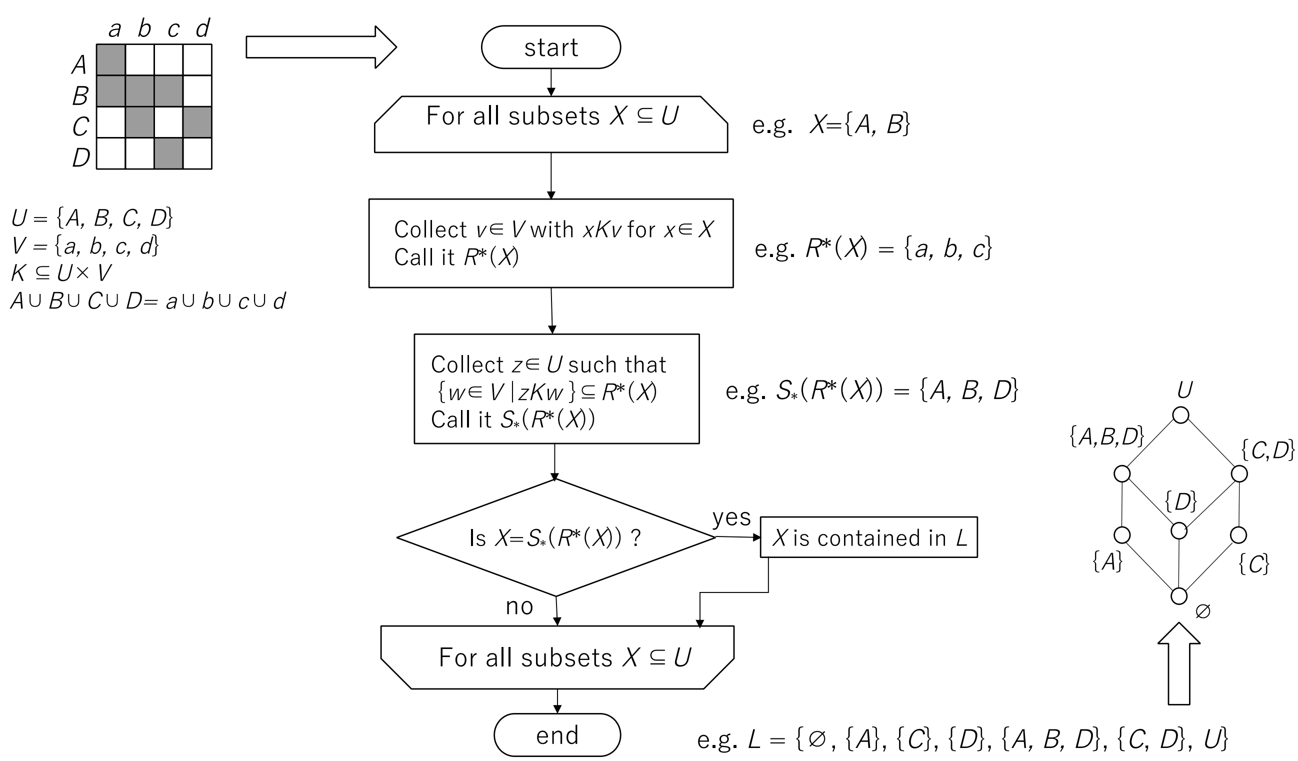

Fig. 6 The flow chart of constructing a rough-set lattice for a given binary relation

divided into disjoint parts. Now, $P(A_i) = Prob([A_i]_L)$ is defined by

$$Prob([A_i]_L) = \sum_{k \in 1[A]_L} kP[A]_L(k)/N \tag{22}$$

where $P_{[A]_L}(x)$ represents the probability of binary sequences in $[A]_L$, assigned by x representing the number of 1 s in the binary sequence, and $\mathbf{1}[A]_L$ represents a set of the number of 1 s in each binary sequence in $[A]_L$. Figure 7 shows the flow chart of calculating the probability of an element of the lattice.

Given $[A]_L = \{01000, 01100, 01010, 01001, 01110\}$, $N = 5$, $\mathbf{1}[A]_L = \{1, 2, 3\}$ and $Prob([A]_L) = (1/5 + 2(3/5) + 3(1/5))/5 = (10/5)/5 = 2/5$ because $P_{[A]_L}(1) = 1/5$, $P_{[A]_L}(2) = 3/5$ and $P_{[A]_L}(3) = 1/5$. Clearly, $Prob([\max L_N]_L) = N/N = 1.0$ and $Prob([\min L_N]_L) \geq 0.0$. In the Boolean and a lattice of almost disjoint union of Boolean algebras, $Prob([\min L_N]_L) = 0.0$ is clear to see. It can be clearly seen that for any A, B in L_N, P is a monotonous map such that

$$A \leq B \Rightarrow P(A) \leq P(B). \tag{23}$$

This can be shown easily by the following:

We can assume that $\max\mathbf{1}[A]_L = n$; then, $\min\mathbf{1}[B]_L = n+1$ without loss of generality. Because of the definition of $[A]_L$ and the assumption $\max\mathbf{1}[A]_L < \min\mathbf{1}[B]_L$, we consider the case in which $[A]_L$ contains the most possible elements, $\mathbf{1}[A]_L = \{1, 2, \ldots, n\}$, without loss of generality. If $k \notin \mathbf{1}[A]_L$, it does not affect the following argument because $P_{[A]L}(k) = 0$. Let $\max\mathbf{1}[B]_L = m$, $\mathbf{1}[B]_L = \{n+1, n+2, \ldots, m\}$. Thus,

$$\begin{aligned} P(A) &= Prob([A]_L) = (P_{[A]_L}(1) + 2P_{[A]_L}(2) + \ldots \\ &\quad + (n-1)P + [A]_L(n-1) + nP_{[A]_L}(n))/N \end{aligned} \tag{24}$$

Because $P_{[A]_L}(1) + P_{[A]_L}(2) + \ldots + P_{[A]_L}(n) = 1$,

$$\begin{aligned} P(A) &= \frac{P_{[A]_L}(1) + 2P_{[A]_L}(2) + \cdots + (n-1)P_{[A]_L}(n-1) + n(1 - P_{[A]_L}(1) - P_{[A]_L}(2) - \cdots - P_{[A]_L}(n-1))}{N} \\ &= \frac{n - ((n-1)P_{[A]_L}(1) + (n-2)P_{[A]_L}(2) + P_{[A]_L}(n-1))}{N} < n/N \end{aligned} \tag{25}$$

$$\begin{aligned} P(B) = &\ ((n+1)P_{[A]_L}(n+1) + (n+2)P_{[A]_L}(n+2) \\ &+ \cdots + (m-1)P_{[A]_L}(m-1) + mP_{[A]_L}(m))/N \\ = &\ ((n+1)(P_{[A]_L}(n+1) + P_{[A]_L}(n+2) + \cdots + P_{[A]_L}(m)) \\ &+ (P_{[A]_L}(n+2) + 2P_{[A]_L}(n+3) + \cdots + (m-n-1)P_{[A]_L}(m)))/N \\ = &\ (n+1)/N + (P_{[A]_L}(n+2) + 2P_{[A]_L}(n+3) + \ldots \\ &+ (m-n-1)P_{[A]_L}(m))/N > (n+1)/N. \end{aligned} \tag{26}$$

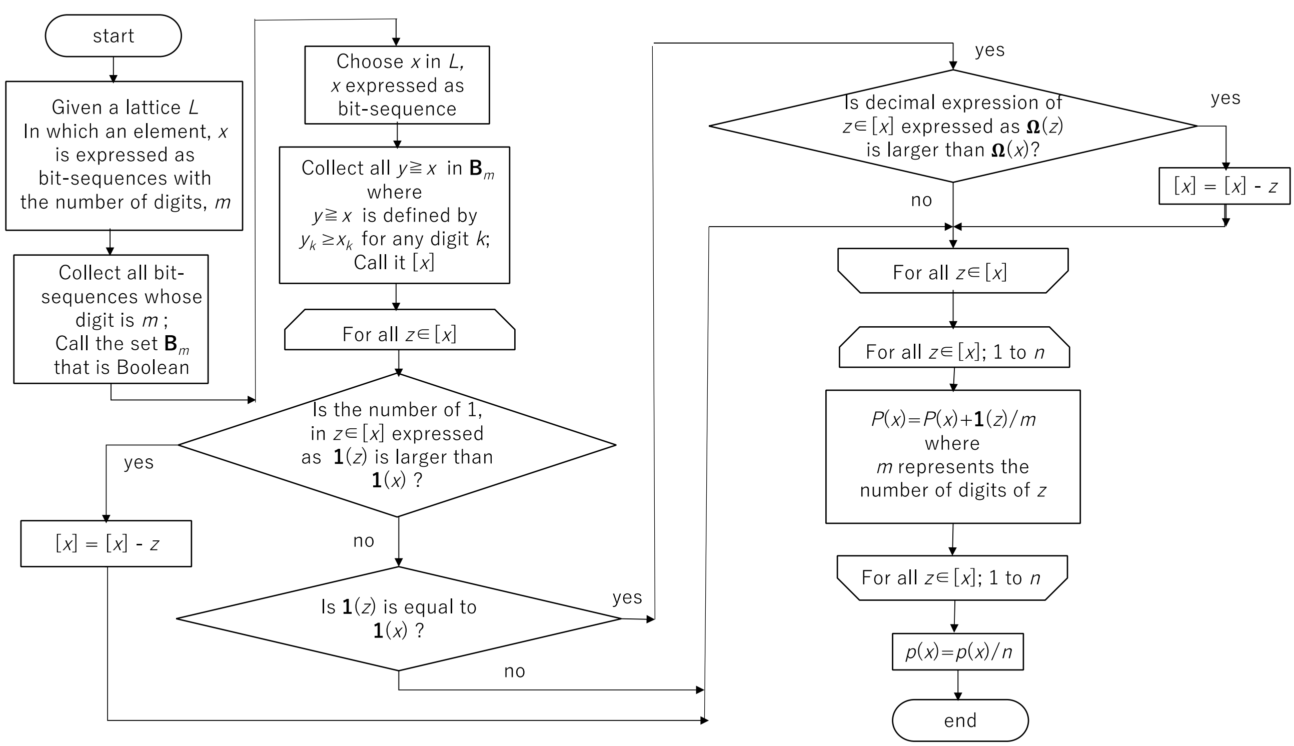

Fig. 7 The flow chart of calculating the probability of an element of the lattice

As a result, $P(A) < n/N < (n+1)/N < P(B)$.

Owing to the monotonousness, $P(A)$, $P(B) \leq P(A \text{ OR } B)$, $P(A \text{ AND } B) \leq P(A)$, and $P(B)$. Thus, we have that

$$P(A \text{ AND } B) \leq (P(A) + P(B))/2 \leq P(A \text{ OR } B). \tag{27}$$

This shows that cognitive illusions such as those associated with the guppy effect and conjunction fallacy cannot be verified in the probability space derived from a single lattice. When we focus on lattice L in which the probability is calculated, we write $P(A)$ as $P_L(A)$. In $\mathbf{B}_N$, $P(A)$ is written as $P_{\mathrm{B}}(A)$, and in $\mathbf{O}_N$, $P(A)$ is written as $P_{\mathrm{O}}(A)$. Similarly, we use $Prob_{\mathrm{B}}(A)$ and $Prob_{\mathrm{O}}(A)$.

In $\mathbf{B}_N$, any $[A]_{\mathrm{B}}$ is a one-element set $\{A\}$, and $P(A) = Prob([A]_{\mathrm{B}}) = Prob(\{A\}) = \mathbf{1}(A)/N$. For any pair of binary sequences A and B, let the number of 1 s occurring in both A and B in the same digit be n_c, the number of 1 s occurring only in A in the same digit be n_A, and the number of 1 s occurring only in B in the same digit be n_B. Thus, $P(A) = (n_A + n_c)/N$, $P(B) = (n_B + n_c)/N$, $P(A \text{ AND } B) = n_c/N$, and $P(A \text{ OR } B) = (n_A + n_B + n_c)/N$. Thus, $P(A) + P(B) - P(A \text{ AND } B) = (n_A + n_B + 2n_c)/N - n_c/N$, and

$$P(A \text{ OR } B) = P(A) + P(B) - P(A \text{ AND } B). \tag{28a}$$

Particularly, it is easy to see that for disjoint sets A and B,

$$P(A \text{ OR } B) = P(A) + P(B). \tag{28b}$$

That implies an additive law. In our framework, the additive law holds only in $\mathbf{B}_N$ and never holds in non-Boolean lattices. Because we focus on the ambiguity of decision-making or incomplete bounded rationality, we give up the law of additivity. However, our framework allows $A \leq B \Rightarrow P(A) \leq P(B)$, which can be considered weakened additivity.

$P(A \text{ AND } B) \leq (P(A) + P(B))/2 \leq P(A \text{ OR } B)$ is also shown by counting the number of 1 s in A and/or B because $(P(A) + P(B))/2 = (n_A + n_B + 2n_c)/2N$. Because $P(A$ and $\text{NOT}(B)) = n_A/N$, it can be seen that

$$P(A) = P(A \text{ AND } B) + P(A \text{ AND NOT } (B)). \tag{29}$$

As mentioned earlier, we assume that decision-making with the logical operations AND, OR, and NOT can modify the probability space from one based on Boolean lattice $\mathbf{B}_N$ to one based on orthomodular lattice $\mathbf{O}_N$.

In our approach, various cognitive illusions can result from the change in probability from a Boolean lattice to an orthomodular lattice. We assume that the probability of single event A (binary sequence) before a logical operation is performed is obtained from $\mathbf{B}_N$. Thus, $P(A)$ is obtained as $P_{\mathrm{B}}(A) = Prob_{\mathrm{B}}([A]_{\mathrm{B}}) = Prob_{\mathrm{B}}(\{A\})$.

In contrast, after a logical operation is performed, $P(A \text{ AND } B)$ is obtained as $P_{\mathrm{O}}(A \text{ AND } B)$, where the AND operation is applied to A and B in $\mathbf{B}_N$ and the probability of A AND B is obtained in $\mathbf{O}_N$. This results in $P(A \text{ AND } B) = Prob_{\mathrm{O}}([A \text{ AND } B]_{\mathrm{O}})$. As mentioned in the next section, this leads to $Prob_{\mathrm{O}}([A \text{ AND } B]_{\mathrm{O}}) > Prob_{\mathrm{B}}([A]_{\mathrm{B}})$, which is similar to the conjunction fallacy.

In our model, multi-contexts which are expressed by multi-dimensional aspects in quantum cognition are expressed by almost disjoint union of multi-Boolean algebras. Thus, practical scenarios can be handled in our model, as well as the quantum psychological model can do [27, 66]. Recently, three-way decisions are proposed, based on the notion of acceptance, rejection, and non-commitment [67], and developed in the relation of quantum formalism [28]. Since our construction contains upper and lower approximations with respect to two binary relations, the combination of them can entail three-way decisions.

Simple Example from Relation to Probability

We show how to construct a lattice and how to calculate probability from a binary relation. Figure 8A shows how to construct a lattice from a diagonal binary relation. A set of attributes $\{a, b, c\}$ and a set of objects $\{A, B, C\}$ is a partition of a universal set, $U = a \cup b \cup c = A \cup B \cup C$, and can be derived from the two kinds of cognitive maps f and g shown in Fig. 2B. A collection of solutions for $X = S_*R^*(X)$, where X is a subset of U. Since it is easy to see that a candidate of a solution to $X = S_*R^*(X)$ is a union of a subset of $\{a, b, c\}$, one must check all subsets of $\{a, b, c\}$ in the table in Fig. 6A.

For example, in taking $X = a \cup b$, which is represented by ab, the elements in $\{A, B, C\}$ having a relation (gray cell) to a or b are collected. Since A has a relation to a and B has a relation to b, one can collect $AB = R^*(ab)$. Then, one calculates $S_*R^*(ab) = S_*(AB)$ and collects the elements in $\{a, b, c\}$ that have no relation to the elements in $\{A, B, C\}$ except for $R^*(ab) = AB$. Thus, one obtains $S_*R^*(ab) = ab$. As shown in the table in Fig. 8A, any unions of elements in $\{a, b, c\}$ can be solutions to $S_*R^*(X) = X$. This implies a power set of $\{a, b, c\}$, which is represented by the Hasse diagram. The number accompanied by an element of a lattice represents the order in a lattice defined by Eqs. (16) and (17). Given the nondiagonal relation shown in Fig. 8B, some subsets are not solutions to $S_*R^*(X) = X$. In taking $X = a$, a collection of elements in $\{A, B, C\}$ having a relation to a is AC since the cells (a, A) and (a, C) are shaded. This implies $R^*(a) = AC$. Then, one collects the elements in $\{a, b, c\}$ having no relation to the elements in $\{A, B, C\}$, except for $R^*(a) = AC$. This is not only a but also c, and one obtains $S_*R^*(a) = ac$. This implies that a is not a solution to $S_*R^*(X) = X$ (not an element of a lattice). An obtained lattice is shown as a Hasse diagram. The number accompanied by an element of a lattice represents the order in a lattice defined by Eqs. (16) and (17).

Figure 8C shows how to calculate the probability of an element of a lattice, L, obtained in Fig. 8B. For element x in a lattice, $[x]_L$ is calculated. As mentioned before, there

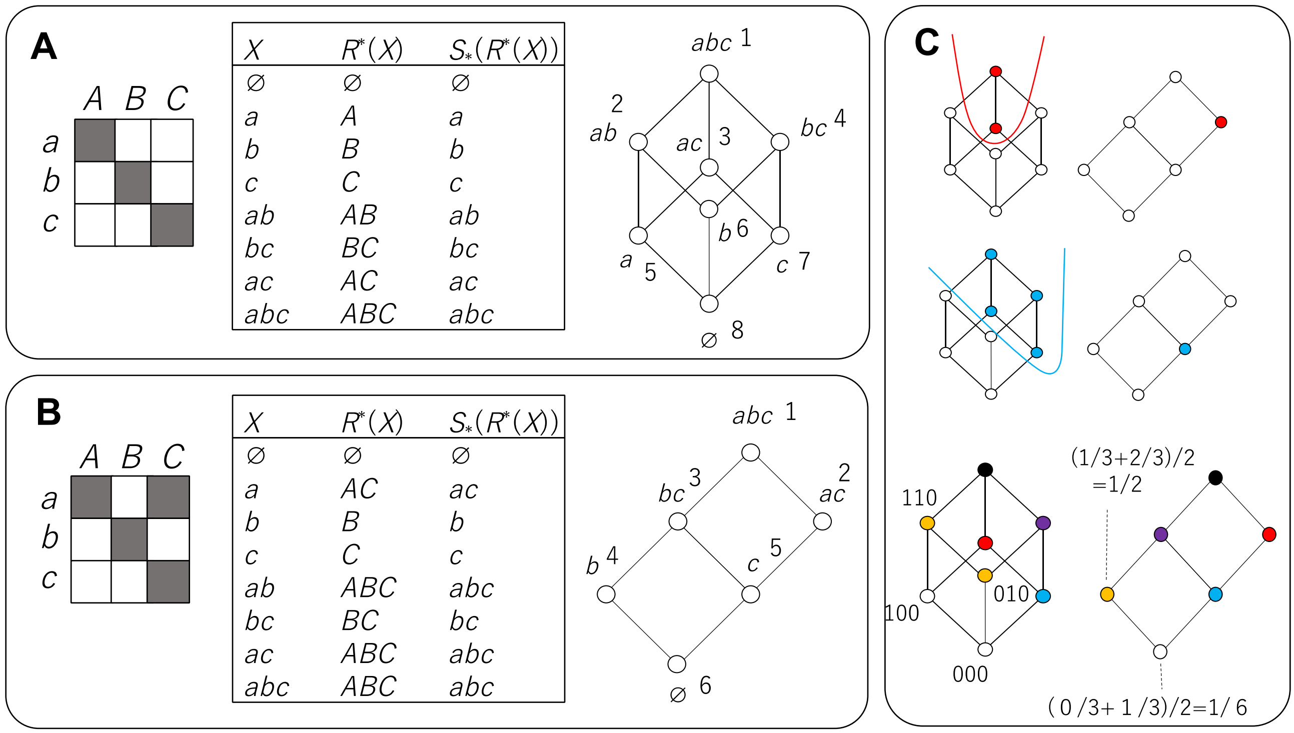

Fig. 8 Binary relation, corresponding lattice, and probability. **A** Given a diagonal relation, any unions of subsets of $\{a, b, c\}$ can be elements of a rough-set lattice. This leads to 2^3-Boolean algebra. **B** Given a nondiagonal relation, some subsets cannot be a solution to $X = S_*R^*(X)$. This leads to a non-Boolean lattice. **C** Probability of an event (binary sequence) in a non-Boolean lattice in **B**

is a Boolean algebra, **B**, containing a given lattice, L. In the case of the lattice shown in Fig. 8B, the corresponding Boolean algebra is the lattice shown in Fig. 8A. Since $[x]_L$ is a disjoint subset of $\mathbf{B}_3$, it is taken by the order of elements of a lattice, L. First, $[abc]$ is taken due to the order since abc is the largest element in $\mathbf{B}_3$, $[abc] = \{abc\}$. Second, $[ac]$ is an upper set of ac except for $[abc]$. One obtains $[ab] = \{x$ in $\mathbf{B}_3 \mid x \geq ab\} - [abc] = \{ab, abc\} - \{abc\} = \{ab\}$. Similarly, $[ac] = \{ac, abc\} - [abc] = \{ac\}$, where the elements of $\{ac, abc\}$ are represented by red. Indeed, $[bc] = \{bc, abc\} - [abc] = \{bc\}$, $[b] = \{b, ab, bc, abc\} - [bc] - [abc] = \{b, ab\}$, and $[c] = \{c, ac, bc, abc\} - [ac] - [bc] - [abc] = \{c\}$, where the elements of $\{c, ac, bc, abc\}$ are represented by blue. Finally, $[\varnothing] = \{\varnothing, a\}$. The probability of element x of a lattice is an average of the probability of the elements in $[x]$. Thus, $P(b) = (P_{\mathbf{B}}(110) + P_{\mathbf{B}}(010))/2 = (2/3 + 1/3)/2 = 1/2$, and $P(\varnothing) = (P_{\mathbf{B}}(100) + P_{\mathbf{B}}(000))/2 = 1/6$

Figure 9 also shows an example of the calculation of the probability of an element of a non-Boolean lattice. A Boolean lattice containing a given lattice is $\mathbf{B}_4$.

Results: Cognitive Illusions

In this section, we show that our model can describe various cognitive illusions that are found in cognitive experiments and that are explained by quantum cognition. Our preliminary explanation can show the possibility of a rigorous explanation of cognitive illusions.

Guppy Effect

Figure 10 shows the change in the probability space from $\mathbf{B}_5$ to $\mathbf{O}_5$. As a subset of $\mathbf{B}_5$, $\{10100, 10101, 10110, 10111\}$ corresponds to an element, 10100, in $\mathbf{O}_5$, $P_{\mathrm{O}}(10100) = (1 \times 4 + 2 \times 3 + 1 \times 2)/4/5 = 3/5 = 0.6$. Thus, we can explain the cause of the guppy effect.

Because a single item is considered in the calculation of $\mathbf{O}_5$ with respect to the probability, $P(11001) = P_{\mathrm{B}}(11001) = 3/5 = 0.6$, and similarly, $P(11010) = P_{\mathrm{B}}(11010) = 0.6$. In contrast, a binary operation modifies the lattice, $P(11001 \text{ AND } 11010) = P_{\mathrm{O}}(11000) = Prob([11000]_{\mathrm{O}}) = Prob(\{11011, 11101, 11110, 11001, 11010, 11100, 11000\}) = (3 \times 4 + 3 \times 3 + 1 \times 2)/7/5 = 23/35 \sim 0.66$. Therefore, when we write $A = 11001$, $B = 11010$,

$$P(A \text{ AND } B) > (P(A) + P(B))/2 \tag{30}$$

This is simply the guppy effect. Similarly, $P(10111) = P_{\mathrm{B}}(10111) = 0.8$, $P(11110) = P_{\mathrm{B}}(11110) = 0.8$, and $P(10111 \text{ AND } 11110) = P_{\mathrm{O}}(10110) = P_{\mathrm{O}}(10100) = 0.6$ because 10110 is included in $[10100]_{\mathrm{O}}$. In this case, the guppy effect does not occur.

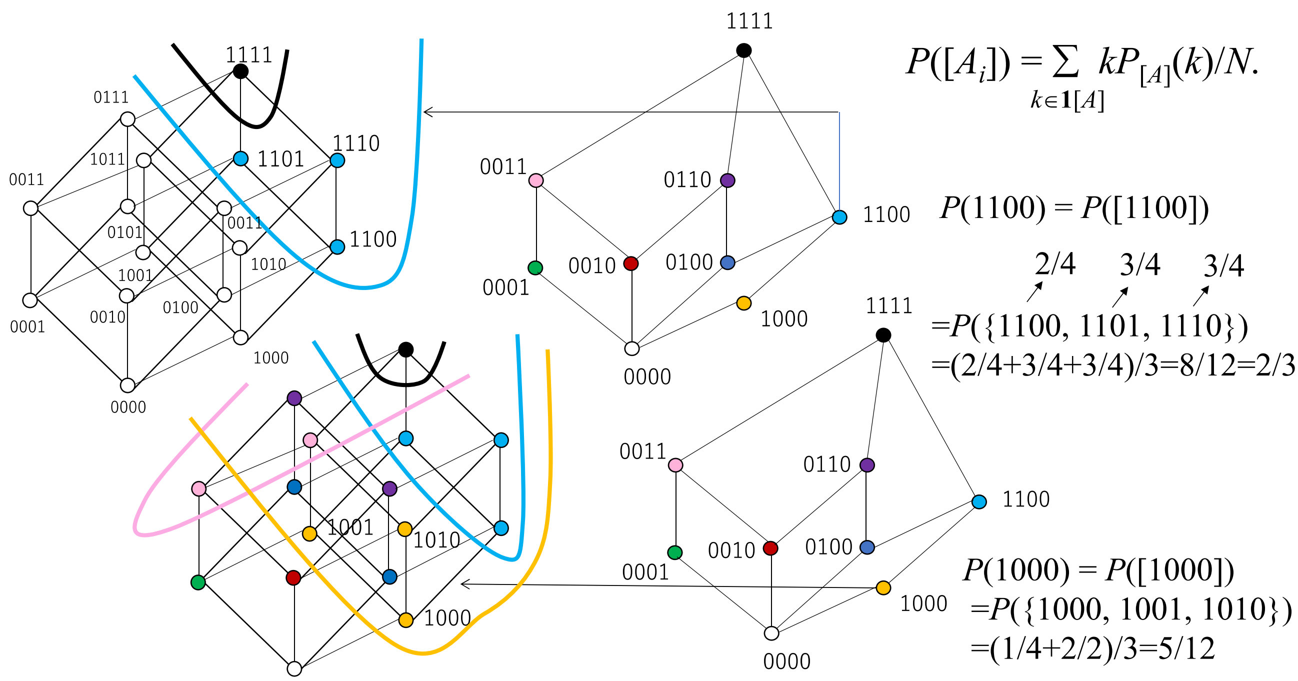

Fig. 9 An example of the calculation of the probability of elements in a non-Boolean lattice. Element x in the non-Boolean lattice and elements of $[x]$ in the $\mathbf{B}_4$-Boolean lattice have the same color

Conjunction Fallacy

The conjunction fallacy also implies that $P(A \text{ and } B) > P(A)$. This is simply an example of the guppy effect. In other words, the conjunction fallacy can be concretely explained in our framework.

Order Effect

The order effect is usually explained using quantum psychology since quantum mechanics can use the order effect in measurement [61]. As mentioned before, quantum psychology assumes bounded rationality. By contrast, we provide an

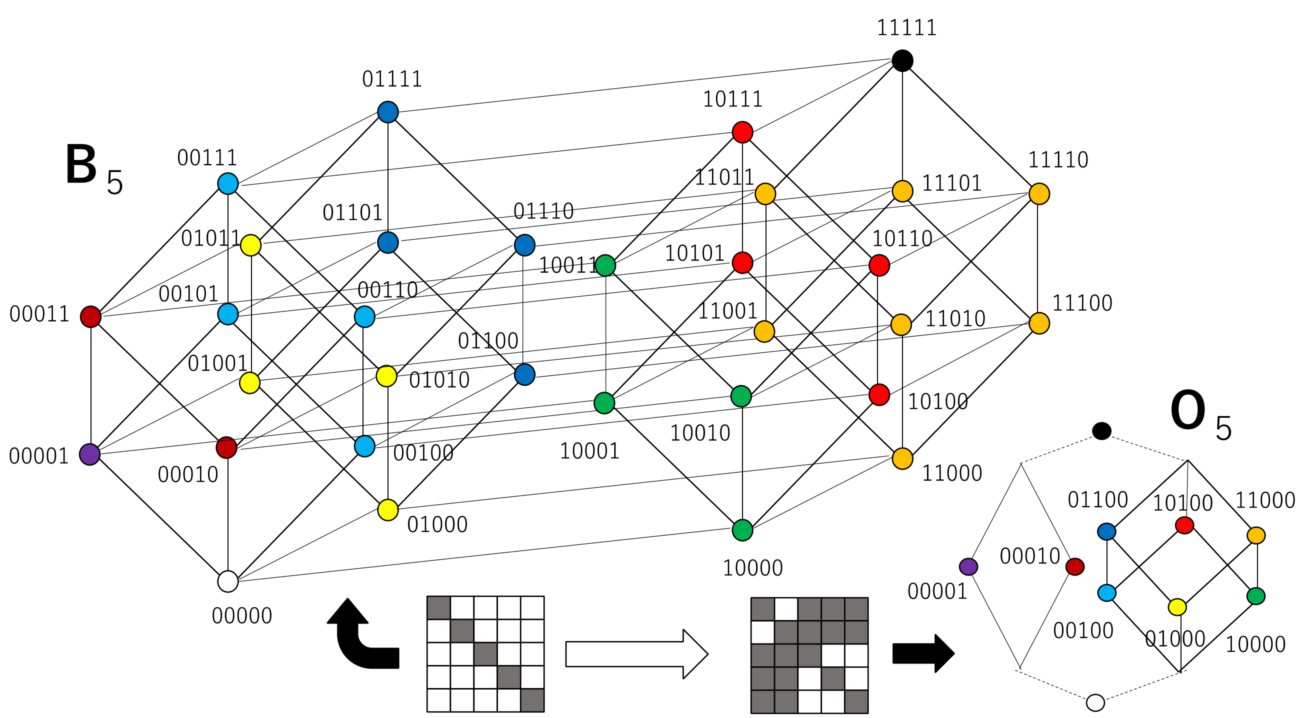

Fig. 10 Change in the probability space from a Boolean lattice, $\mathbf{B}_5$, to an orthomodular lattice, $\mathbf{O}_5$. Each element of $\mathbf{O}_5$ is expressed as a subset of $\mathbf{B}_5$; the elements of a subset of $\mathbf{B}_5$ corresponding to an element of $\mathbf{O}_5$ have the same color

explanation under our framework with incomplete bounded rationality.

The order effect, $P(A \text{ AND } B) \neq P(B \text{ AND } A)$, is easily explained by our model. Let categories A and B be expressed as 2 by 2 and 3 by 3 diagonal relations, respectively (Fig. 11). The order A and B divides a 5 by 5 binary relation into a 2 by 2 relation and a 3 by 3 relation from left to right, and the order B and A divides a 5 by 5 binary relation into a 3 by 3 relation and a 2 by 2 relation. As atoms are numbered using a binary relation from left to right, 00001, 00010, …, 10000, partitions are changed depending on the order of A and B.

If the AND operation is applied to A before B (Fig. 4 above), the atoms of the $\mathbf{B}_2$ lattice are 00001 and 00010, and those of the $\mathbf{B}_3$ lattice are 00100, 01000, and 10000. In this case, for example, $[00100]_{\text{O}} = \{00100, 00101, 00110, 00111\}$ and $[10000]_{\text{O}} = \{10000, 10001, 10010, 10011\}$. Thus, $P_{\text{O}}(00100) = P_{\text{O}}(10000) = (3 \times 1 + 2 \times 2 + 1 \times 1)/4/5 = 0.4$. In contrast, if the AND operation is applied to B before A (Fig. 9), the atoms of the $\mathbf{B}_3$ lattice are 00001, 00010, and 00100, and those of the $\mathbf{B}_2$ lattice are 01000 and 10000. In this case, $[00100]_{\text{O}} = \{00100\}$ and $P_{\text{O}}(00,100) = (1 \times 1)/1/5 = 0.2$. Similarly, $[10000]_{\text{O}} = \{10000, 10001, 10010, 10100, 11000, 11001, 11010, 11100\}$ and $P_{\text{O}}(10000) = (3 \times 3 + 4 \times 2 + 1 \times 1)/8/5 = 0.45$. Note that 10000 = 10001 AND 11000. It is clear that 10001 is above 00001; however, 11000 is not. In this sense, 10001 belongs to A, and 11000 belongs to B. Thus, the order A before B refers to 10001 AND 11000, and the order B and A refers to 11000 AND 10001. Because the order affects the structure of the lattice and probability space, we obtain $P_{\text{O}}(10001 \text{ AND } 11000) = 0.4 < 0.45 = P_{\text{O}}(11000 \text{ AND } 10001)$. This implies that there exist A and B in $\mathbf{B}_N$ such that $P(A \text{ AND } B) \neq P(B \text{ AND } A)$. This is simply the order effect.

We can describe the order effect without quantum theory as shown in Fig. 11. Although a binary relation consisting of 2 by 2 context and 3 by 3 context is the same as a binary relation consisting of 3 by 3 context and 2 by 2, the numbering elements of corresponding lattices are different from each other. Thus, the probability of elements of the two quasi-disjoint of Boolean algebras are different from each other, and that can explain to the order effect.

Ellsberg and Machina Paradoxes

These paradoxes involve not only the AND operation but also the NOT operation and are expressed as $P(A) \neq P(A \text{ AND } B) + P(A \text{ AND NOT}(B))$. It is clear that $P(A) = P(A \text{ AND } B) + P(A \text{ AND NOT}(B))$ in classical probability theory. As mentioned earlier, both AND and NOT operations are defined in Boolean algebra in our approach, whereas the elements resulting from the AND operation are interpreted in the orthomodular lattice. Thus, the probability of the elements resulting

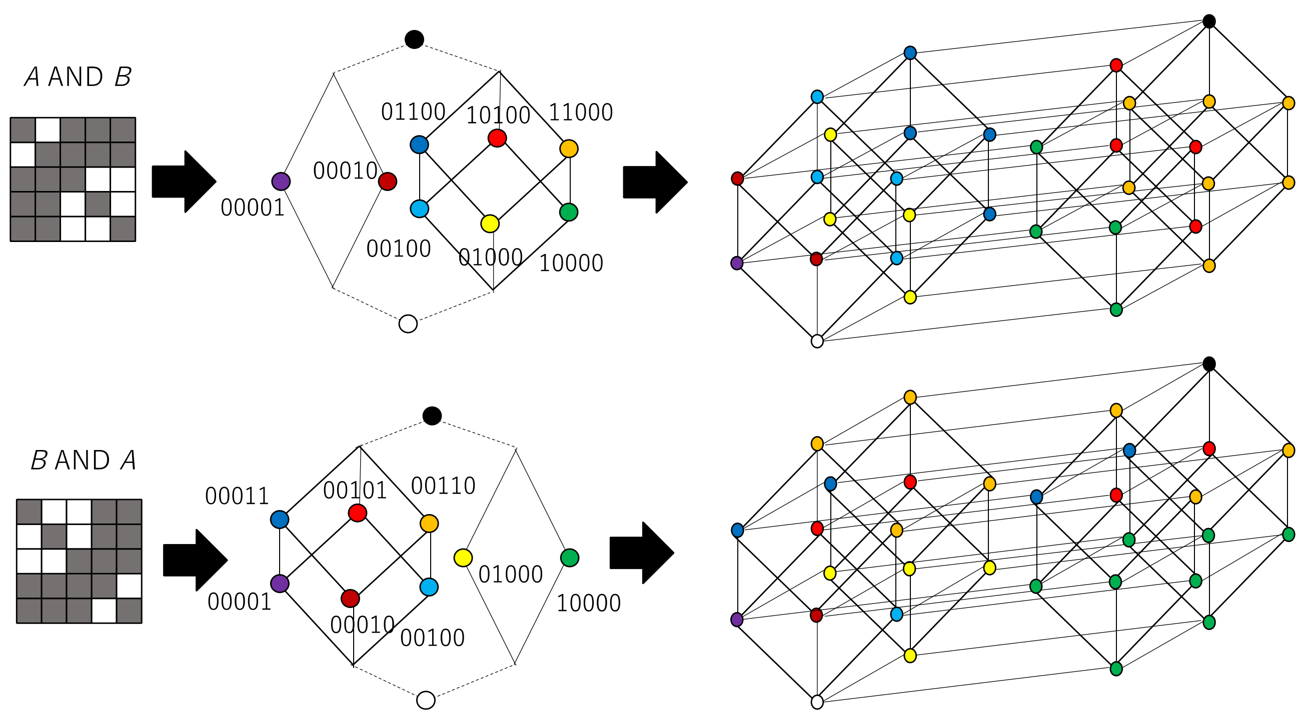

Fig. 11 Order effect explained by the order of sublattices sensitive to a change in the probability space. The order A and B refers to the order of the sublattices from left to right, $\mathbf{B}_2$ and $\mathbf{B}_3$, and the order B and A refers to the order $\mathbf{B}_3$ and $\mathbf{B}_2$ in the orthomodular lattice (center). The elements of the orthomodular lattice that constitute a subset of $\mathbf{B}_5$ (right) are indicated with the same color

from the AND operation is obtained by the map from the orthomodular lattice to [0, 1]. Now, we assume that $A = 11001$ and $B = 11010$ in the case shown in Fig. 3. From the definition of the probability of a single item, $P(11001) = P_B(11001) = 3/5 = 0.6$. In contrast, $P(11001 \text{ AND } 11010) = P_O(11000) \sim 0.66$, and $P(11001 \text{ AND NOT}(11010)) = P(11001 \text{ AND } 00101) = P_O(00001) = 1/5 = 0.2$. As a result, $P(11001) \neq P(11001 \text{ AND } 11010) + P(11001 \text{ AND NOT}(11010))$. This implies the Ellsberg and Machina paradoxes.

Borderline Contradiction

Borderline contradiction is explained using quantum psychology [58]. It can utilize joint probability in quantum cognition and violates classical probability theory. Our description can provide an explanation under quantum-like probability resulting from incomplete bounded rationality.

In probability theory, the law of noncontradiction, NOT(A AND NOT(A)), is expressed as $P(A \text{ AND NOT}(A)) = 0$. The borderline contradiction found in cognitive processes contradicts this law and is expressed as $P(A \text{ AND NOT}(A)) > 0$. Borderline contradiction can be explained by switching between $\mathbf{B}_N$ and $\mathbf{O}_N$ in our approach. This assumes that a change in the logical structure is caused by unstable decision-making between $\mathbf{B}_N$ and $\mathbf{O}_N$. Borderline contradiction is expressed as the back-and-forth transformation between an element of $\mathbf{B}_N$ and a subset of $\mathbf{B}_N$. For any A in $\mathbf{B}_N$, $[A]_O$ is calculated, and for any subset of $\mathbf{B}_N$, S, $\text{rev}[S]_B$ is an element chosen from S.

We now demonstrate borderline contradiction in $\mathbf{O}_5$, as shown in Fig. 3. First, we select $A = 01010$ in $\mathbf{B}_5$; the logical operation NOT is also applied to A in $\mathbf{B}_5$. This results in $\text{NOT}(A) = 10101$. Next, we apply $[\bullet]_O$ to A and NOT(A) and then apply $\text{rev}[\bullet]_B$. Thus, $[A]_O = [01010]_O = \{01011, 01001, 01010, 01000\}$ and $[\text{NOT}(A)]_O = [10101] = \{10111, 10101, 10110, 10100\}$. Finally, $\text{rev}[[A]_O]_B = 01001$ and $\text{rev}[[\text{NOT}(A)]_O]_B = 10101$ are possible. In this case, $P(\text{rev}[[A]_O]_B \text{ AND rev}[[\text{NOT}(A)]_O]_B) = P(10101 \text{ AND } 01001) = P(00001) = 0.2 > 0$. This implies that there is an A in $\mathbf{B}_N$ such that $P(A \text{ AND NOT}(A)) \sim P(\text{rev}[[A]_O]_B \text{ AND rev}[[\text{NOT}(A)]_O]_B) > 0$. In our approach, borderline contradiction results from instabilities in the logical structure. We describe various cognitive illusion by using our probability theory based on the lattice derived from cognitive model featured with nonlocality in cognition.

Finally, we here discuss the significance of our quantum-like probability based on nonlocality in cognition. Many researchers consider that quantum cognition can describe not only cognitive illusion but the essential property of cognitive concept that is phenomenal entity [68]. Conceptual thinking, decision-making, and/or cognition are experienced by a consciousness as a subject [69]. Since consciousness as a subject can be mixed with up a phenomenon in front of the entity, consciousness as a subject can experience a phenomenon. That nature is strongly related to quantum theory on one hand [70], and experiencing subject is regarded as phenomenal entity called phenomenal consciousness [71], and bridging brain data with theoretical studies is proceeded [72]. Quantum theory is expected to describe phenomenal consciousness.

As mentioned in "Motivation of Incomplete Bounded Rationality," we point out the weakness in the bounded rationality, that is the well-known foundation of quantum theory in cognitive process, and expand such an idea to the incomplete bounded rationality. While we give up direct relation between cognitive process and quantum theory, the resulting model in the form of a lattice, that is almost disjoint union of Boolean algebras, can be regarded as a weaken quantum logic in the strict sense of Atmanspacher [29, 31]. Therefore, our approach is consistent with the underlying idea in quantum theory. As mentioned in "Cognitive Illusions and Concept Formation" and "Methods," such a weaken quantum logic is derived from a specific binary relation between objects and attributes, that is called nonlocality in cognition. While the binary relation of the nonlocality in cognition is expressed in very restricted form, the nonlocality reveals that each object has one to one relation between objects and attributes and has identity of itself, while the object is related to attributes outside the context. In other words, the nonlocality in cognition is an entity which can be expanded toward the outside of the context. That is nothing but a restricted expression for phenomenal entity related to phenomenal consciousness. Therefore, we also say that our approach night describes not only cognitive illusion but phenomenal consciousness.

While quantum cognition is recently connected with Bayesian inference [73], quantum-like cognitive process is strongly connected with Bayesian and inverse Bayesian process [63]. Cognitive process is recently found not only in human cognition but also in various living systems including unicellular organisms, physarum [74], and bio-chemical reaction networks [75]. Underlying mechanism in proto-cognitive system is studied in the context of inference system, especially Bayesian inference. Recently, use of Bayesian inference is verified by free energy principle [76], and active modification of environment is verified by active inference [77]. It is known that Bayesian inference is expressed as Boolean algebra or classical logic [63]. While the probability of hypothesis is replaced by the joint probability of the hypothesis under some data, the likelihood of the hypothesis is invariant in Bayesian inference. However, recent studies of which living systems are interpreted as the inference system find that the likelihood of hypothesis is perpetually modified and empirical data are canceled, which is called inverse Bayesian inference [63, 78]. These modifications and canceling avoid a system from steady-state and then make a system wander from a context to another context [78]. Therefore, such inference immediately entails the inference under multi-contexts. Since multi-contexts are expressed by the binary relation with the non-locality of cognition in our model, the corresponding lattice is represented by an almost disjoint union of Boolean lattice or an orthomodular lattice. It results in that quantum-like cognition might be found in various levels of living systems.

Logical operation in orthomodular lattice might entail scale-free properties exhibited by the brain [79]. Since orthomodular lattice is represented by multiple Boolean sub-lattices whose greatest and least elements are common, decision-making is realized in one Boolean sub-lattices and wanders from one sub-lattice to another sub-lattice [63, 78]. So far as sub-lattices are regarded as attracters in a dynamical system of the brain, wandering process is chaotic wandering from one attracter to another attracter. It is recently reported that an agent with both Bayesian and inverse Bayesian inference shows scale-free properties [80, 81]. As mentioned before, the inference consisting of Bayesian and inverse Bayesian inference can entail orthomodular lattice [63]. Thus, inference process in orthomodular lattice might underlie the mechanism of scale-free properties in the brain.

Conclusion

Although concept formation and cognitive illusions appear to share a mechanism, they have been investigated independently. While quantum cognition can successfully explain various cognitive illusions, the reason for the introduction of quantum mechanics is based on axiomatic bounded rationality. Since the fundamental assumption in bounded rationality is still unclear, we expand the idea of bounded rationality and introduce incomplete bounded rationality. Under this condition, we do not verify the foundation of quantum theory. Rather, we define the new cognitive model without quantum theory, and that is based on the incomplete bounded rationality. Indeed, our model looks quantum-like one because our mode is expressed as a quasi-disjoint system of Boolean algebras containing an orthomodular lattice which represents quantum logic in the form of a lattice.

In this paper, we introduce nonlocality in cognition, where an object A is identified as A rather than not A. If the nonlocality is expressed as a relation between objects and attributes, both concept formation and cognitive illusions can be explained using our framework. Using a rough-set lattice technique, concepts are expressed as elements of a lattice derived from the relation. By defining a probability in the lattice, cognitive illusions can be described through the change in the probability space triggered by actual decision-making. This provides a full explanation of cognitive illusions under incomplete bounded rationality without quantum theory, while the model results in quantum-like model or the model featuring weak quantum.

As nonlocality in cognition can require an orthomodular lattice corresponding to quantum mechanics, the quantum mechanical-like mathematical structure is verified to originate in a general cognitive world without quantum mechanics itself. Our idea can be a natural extension of quantum mechanical thinking.

Appendix 1. Lattice Theory

Here, we define an ordered set by a collection of sets, where an order is defined by inclusion. The binary operations OR and AND are defined as

$$A \text{ OR } B = \min\{X \in L | A \subseteq X, B \subseteq X\}$$

$$A \text{ AND } B = \max\{X \in L | X \subseteq A, X \subseteq B\}.$$

If an ordered set P is closed with respect to OR and AND, P is called a lattice. A lattice has the smallest and largest elements, which are represented by 0 and 1, respectively. The complementary NOT operation is defined as follows. For any A in a lattice, NOT(A) is an element of a lattice such that

$$A \text{ AND NOT } (A) = 0, \quad A \text{ OR NOT } (A) = 1.$$

Figure 12 shows Hasse diagrams of some lattices obtained from a collection of subsets of $\{a, b, c\}$. In the Hasse diagram, all elements of a lattice are represented by circles. If element A is smaller than element B (i.e., A is included in B), and there is no other element between A and B, the two elements are linked by a line, and A is located below B. Figure 12A shows a lattice obtained from all subsets of $\{a, b, c\}$, where OR and AND are expressed as a union ($\cup$) and an intersection ($\cap$), respectively. This lattice is a Boolean lattice. In contrast, Fig. 12C shows that OR is not expressed as a union, where $\{a\}\text{OR}\{b\} = \{a, b, c\} \neq \{a, b\} = \{a\} \cup \{b\}$.

Especially, in a lattice theory, the nearest element being larger than the least element is called an atom, which are not related to physical atom.

Appendix 2. Rough-Set Lattice

Figure 13 shows an example of the procedure for obtaining a lattice from a binary relation. From a binary relation between objects and attributes, one can see that one object has multiple attributes such that object A has attributes a and d, as seen in the left part of the figure. An object is considered to consist of virtual elements x_1 and x_2, which can be interpreted as attributes a and d, respectively. All virtual

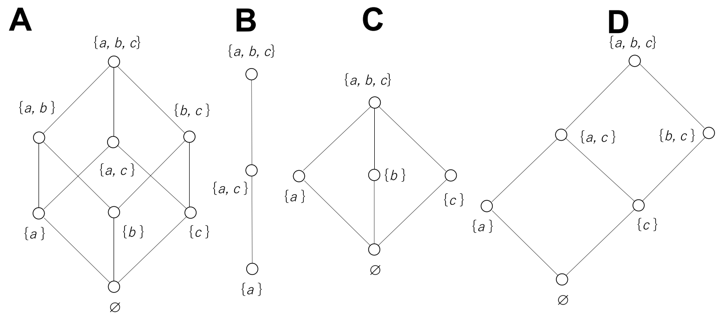

Fig. 12 A–D Hasse diagrams of some lattices obtained from a family of subsets of $\{a, b, c\}$

elements contained in objects are numbered in the order $x_1, x_2, \ldots, x_s$, where s is the number of related elements in a binary relation. As shown in the center of Fig. 13, A consists of x_1 and x_2; B consists of x_3 and x_4; C consists of x_5, x_6, and x_7; D consists of x_8 and x_9; and E consists of x_{10}, x_{11}, and x_{12}. In other words, a universal set $U = \{x_1, x_2, \ldots, x_{12}\}$ is partitioned in two ways: $A, B, \ldots, E$ and $a, b, \ldots, d$. A partition is derived from an equivalence relation.

Here, we define equivalence relations R and S, which entail the partitions $A, B, \ldots, E$ and $a, b, \ldots, d$, respectively. Object C contains x_5, x_6, and x_7 because x_5, x_6, and x_7 are equal to one another in the sense of equivalence relation R. This relation is written here as x_5Rx_6 and x_6Rx_7 (which also implies x_5Rx_7). Similarly, the equivalence relation S defines x_4Rx_9 and x_9Rx_{11}. Each object is defined as an equivalence class, $[y]_R = \{x \in U | \, xRy\}$, and each attribute is defined by $[y]_S = \{x \in U \mid xSy\}$. Using these equivalence classes, one can define two types of rough sets for set X included in U, such as $R_*(X)$ and $R^*(X)$.

In our framework, there are two types of equivalence relations leading to objects and attributes, R and S. As it is impossible to choose either R or S as a necessary and sufficient condition, we use one (R or S) as a sufficient condition and the other as a necessary condition. A collection of X such that $R_*(S^*(X)) = X$ is a lattice. Set X satisfying this equation is a concept in a rough-set lattice. Any lattice can be obtained in the form of

$$L = \{X \subseteq U | S * (R^*(X)) = X\},$$

if adequate relations R and S are given. In the case of Fig. 13 (left), as $S_*(R^*(a)) = S_*(A \cup C \cup D \cup E)) = a \cup d$, a is not an element of L. It is easy to see that $S_*(R^*(b)) = b$, $S_*(R^*(c)) = c$ and $S_*(R^*(d)) = d$, and atomic elements b, c, and d are elements of L. For two-element sets, $S_*(R^*(b \cup c))) = S_*(B \cup C \cup D \cup$

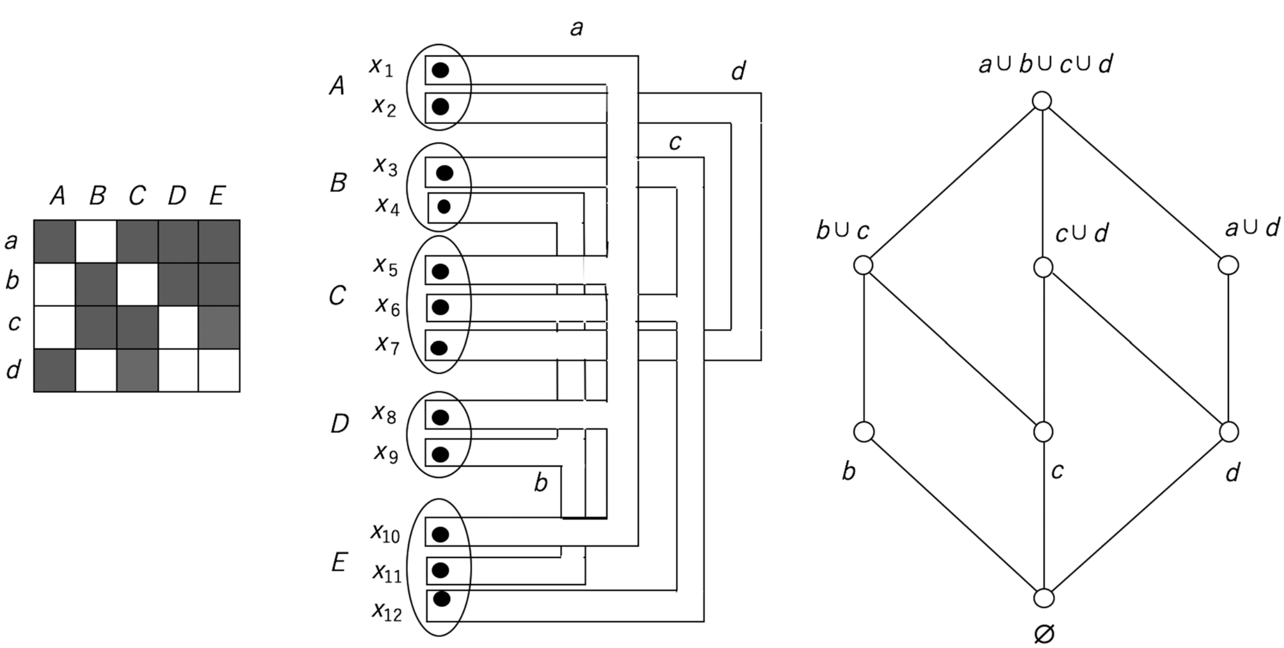

Fig. 13 Construction of a rough-set lattice. A binary relation between objects (columns) and attributes (rows) (left). A box at (x, Y) is shaded if x is related to Y. Two types of partitions for a set of virtual elements (center). Hasse diagram of a rough-set lattice obtained from the binary relation (right)

$E) = b \cup c$, $S_*(R^*(c \cup d)) = S_*(A \cup B \cup C \cup E) = c \cup d$ and $S_*(R^*(a \cup d)) = S_*(A \cup C \cup D \cup E) = a \cup d$; however, no other two-element unions include elements of L. In all three element unions, X_3 is not an element of L, as $S_*(R^*(X_3)) = S_*(U) = U$, where $U = a \cup b \cup c \cup d = A \cup B \cup C \cup D \cup E$. Finally, we obtain $S_*(R^*(U)) = U$ and $S_*(R^*(\varnothing)) = \varnothing$. The right part of Fig. 13 shows a Hasse diagram of $L = \{X \subseteq U | S_*(R^*(X)) = X\}$ obtained from a binary relation shown in the left of the figure.

Funding This work was funded by JSPS (Japan Society of the Promotion of Science) Kaken-hi 18K18478 and JPJS 00120351748.

Declarations

Informed Consent This article does not contain any experiments required for informed consent.

Human and Animal Rights This article does not contain any studies involving human participants and/or animals by any of the authors.

Conflict of Interest The authors declare no competing interests.

References

1. Gilovich T, Griffin D, Kahneman D. Heuristics and biases: the psychology of intuitive judgment. Cambridge University Press; 2002.
2. Rosch E. Natural categories. Cogn Psychol. 1973;4:328–50.
3. Kahneman D. Judgment under uncertainty: heuristics and biases. Cambridge University Press; 1982.
4. Ganter B, Wille R. F*ormal concept analysis: mathematical foundations*. Springer; 1999.
5. Ganter B, Meschke C. A formal concept analysis approach to rough data tables. In Rough sets, fuzzy sets, data mining and granular computing (pp. 117–126). Heidelberg: Springer Berlin; 2009.
6. Pawlak Z. Information systems-theoretical foundations. Inf Syst. 1981;6:205–18.
7. Polkowski L. Rough sets, mathematical foundations. Heidelberg: Springer; 2002.
8. Yao YY. A comparative study of formal concept analysis and rough set theory in data analysis. In: Tsumoto S, Slowinski R, Komorowski J, Grzymala-Busse JW, (eds.)RSCTC 2004. LNCS (LNAI), 3066, (pp. 59–68). Heidelberg: Springer; 2004.
9. Järvinen J. Pawlak's information systems in terms of Galois connections and functional dependencies. Fundam Inform. 2007;75:315–30.
10. Gunji Y-P, Haruna T. A non-Boolean lattice derived by double indiscernibility. Transactions on Rough Sets XII. 2010;211–225.
11. Chater N, Tenenbaum JB, Yuille A. Probabilistic models of cognition: conceptual foundations. Trends Cogn Sci. 2006;10:287–91.
12. Tenenbaum JB, Griffiths TL, Kemp C. Theory-based Bayesian models of inductive learning and reasoning. Trends Cogn Sci. 2006;10:309–18.
13. Manktelow K. Thinking and reasoning: an introduction to the psychology of reason. Judgment and Decision Making: London, Psychology Press; 2012.
14. Gudder SP. Quantum probability. Academic Press; 1988.
15. Aerts D. Quantum structure in cognition. J Math Psychol. 2009;53:314–48.
16. Aerts D, Gabora L, Sozzo S. Concepts and their dynamics: A quantum-theoretic modeling of human thought. Top Cogn Sci. 2013;5:737–72.
17. Busemeyer JR, Bruza PD. Quantum models of cognition and decision. Cambridge University Press; 2012.
18. Khrennikov A. Linear representations of probabilistic transformations induced by context transitions. J Phys A. 2001;34:9965–81.
19. Osherson D, Smith E. On the adequacy of prototype theory as a theory of concepts. Cognition. 1981;9:35–58.
20. Hampton J. Testing the prototype theory of concepts. J Mem Lang. 1995;34:686–708.
21. Alxatib S, Pelletier J. On the psychology of truth gaps. In: Nouwen R, van Rooij R, Sauer- land U, Schmitz H-C, editors. Vagueness in Communication. Berlin, Heidelberg: Springer-Verlag; 2011. p. 13–36.
22. Machina MJ. Risk, ambiguity, and the dark-dependence axioms. Am Econ Rev. 2009;99:385–92.
23. Bruza PD, Wang Z, Busemeyer JR. Quantum cognition: a new theoretical approach to psychology. Trends Cogn Sci. 2015;19:383–93.
24. Aerts D, Broekaert J, Gabora L, Veroz T. The guppy effect as interference. Quantum Interaction 2012. Springer; 2012. p.36–74.
25. Busemeyer JR, Franco R, Pothos EM, Franco R, Trueblood JS. A quantum theoretical explanation for probability judgment errors. Psychol Rev. 2011;118(2):193–218.
26. Blutner R, beim Graben, R. Quantum cognition and bounded rationality. Synthese. 2016;193:3239–91.
27. Ishwarya MS, Kumar ChA. Quantum aspects of high dimensional formal representation of conceptual spaces: a model for achieving consciousness. Cogn Comput. 2020;12:563–76.
28. Ishwarya MS, Kumar ChA. Decision-making in cognitive paradoxes with contextuality and quantum formalism. Appl Soft Comput. 2020;95:106521.
29. Narens L. Probabilistic lattices: with applications to psychology. Singapore: World Scientific; 2014.
30. Svozil K. Randomness and undecidability in Physics. Singapore: World Scientific; 1993.
31. Atmanspacher H, Römer H, Walach H. Weak quantum theory: complementarity and entanglement in physics and beyond. Found Phys. 2002;32(3):379–406.
32. Wilce A. Quantum logic and probability theory. In: Zalta EN, editor. The Stanford encyclopedia of philosophy. Stanford: CSLI; 2012.
33. beim Graben P, Atmanspacher H. Extending the philosophical significance of the idea of complementarity. In: Atmanspacher H, Primas H, editors. Recasting reality. Wolfgang Pauli's philosophical ideas and contemporary science. Berlin: Springer; 2009. p. 99–113.

34. Axler S. Linear algebra done right. New York: Springer; 1996.
35. Foulis DJ. A half century of quantum logic-what have we learned?. In: Aerts D, Pykacz J, editors. Quantum structures and the nature of reality : The Indigo Book of Einstein meets Magritte. Dordrecht: Kluwer; 1999.
36. Heunen C. An embedding theorem for Hilbert categories. Theory Appl Categ. 2009;22:321–44.
37. Abramsky S. Contextual semantics; from quantum mechanics to logic, databases, constraints, and complexity. In: Dzhafarov E, et al., editors. Contextuality from Quantum Physics to Psychology, World Scientific; 2015. p.23–50.
38. Manktelow K, Over D, Elqayam S. The science of reason. A Festschrift for Jonathan St B.T. Evans: Psychology Press; 2015.
39. Wickens CD, Hollands JG. Engineering psychology and human performance. Prentice-Hall Inc; 2015.
40. Lakoff G. Women, fire, and dangerous things. What categories reveal about the mind: Oxford University Press; 1987.
41. Pohl RF. Cognitive illusions, intriguing phenomena in thinking, judgement and memory. London: Routledge; 2017.
42. Hampton J. Inheritance of attributes in natural concept conjunctions. Mem Cogn. 1987;15:55–71.
43. Tversky A, Kahneman D. Judgments of and by representativeness. In Kahneman D, Slovic P, Tversky A, editors. Judgment under uncertainty: heuristics and biases. Cambridge, UK: Cambridge University Press; 1982.
44. Schuman H, Presser S. Questions and answers in attitude surveys: 16 experiments on question form, wording, and context. Academic Press; 1981.
45. Peter B, Zheng W, Busemeyer JR. Quantum cognition: a new theoretical approach to psychology. Trends Cogn Sci. 2015;19(7):383–93.
46. Ellsberg D. Risk, ambiguity, and the savage axioms. Q J Econ. 1961;75:643–69.
47. Aerts D, Geriente S, Moreira C, Sozzo S. Testing ambiguity and Machina preferences within a quantum-theoretic framework for decision-making. J Math Econ. 2018;78:176–85.
48. Aerts D, Sozzo S, Tapia J. A quantum model for the Ellsberg and Machina paradoxes. LNCS. 2012;7620:48–59.
49. Bonini N, Osherson DN, Viale R, Williamson T. On the psychology of vague predicates. Mind Lang. 1999;14:377–93.
50. Aerts D, Sozzo S, Veloz T. New fundamental evidence of non-classical structure in the combination of natural concept. Phil Trans R Soc A. 2015;374:20150095.
51. Alxatib S, Pelletier J. On the psychology of truth-gaps. In: Nouwen R, van Rooij R, Sauerland U, Schmitz H-C, editors. Vagueness in Communication. Berlin, Heidelberg: Springer-Verlag; 2011. p. 13–36.
52. Khrennikov AY. Ubiquitous quantum structure. Berlin, Germany: Springer; 2010.
53. Aerts D, Sozzo S. Quantum structure in cognition. Why and how concepts are entangled. In: Song D, Melucci M, Frommholz I, Zhang P, Wang L, Arafat S, editors. Quantum interaction. Berlin, Germany: Springer; 2011. p. 116–27.
54. Rosch E. Prototype classification and logical classification: the two systems. In: Scholnick EK, editor. New trends in conceptual representation: Challenges to Piaget theory? Lawrence Erlbaum: New Jersey; 1983. p. 133–59.
55. Aerts D, Gabora L. A theory of concepts and their combinations I, The structure of the sets of contexts and properties. Kybernetics. 2005;34(1/2):167–91.
56. Aerts D, Aerts S, Broelkaert J, Gabora L. The violation of Bell inequalities n the macroworld. Found Phys. 2000;30(9):1387–414.
57. Baars BJ. A cognitive theory of consciousness. Cambridge, MA: Cambridge Univ. Press; 1989.
58. Baars BJ. Global workspace theory of consciousness; toward a cognitive neuroscience of human experience. Prog Brain Res. 2005;150:454–554.
59. Dehaene S, Kerszberg M, Changeux J-P. A neuronal model of a global workspace in effortful cognitive tasks. Proc Natl Acad Scie USA. 1998;95:14529–34.
60. Gunji Y-P, Sonoda K, Basios V. Quantum cognition based on an ambiguous representation derived from a rough set approximation. BioSystems. 2016;141:55–66.
61. Wang Z, Busemeyer JR. A quantum question order model supported by empirical tests of an a priori and precise prediction. Top Cogn Sci. 2013;5(4):689–710.
62. Blutner R, Pothos EM, Bruza P. A quantum probability perspective on borderline vagueness. Topics in Cognitive Sciences. 2013;5:711–36.
63. Gunji YP, Shinohara S, HarunaT BV. Inverse Bayesian inference as a key of consciousness featuring a macroscopic quantum logic structure. BioSystems. 2017;152:44–63.
64. Gunji YP, Nakamura K. Dancing chief in a brain or consciousness as entanglement. Found Sci. 2020;25:151–84.
65. Gunji YP, Nakamura K, Minoura M, Adamatzky A. Three types of logical structure resulting from the trilemma of free will, determinism and locality. BioSystems. 2020;195: 104151. https://doi.org/10.1016/j.biosystems.2020.104151.
66. Ishwarya MS, Kumar CA. Quantum inspired high dimensional conceptual space as KID model for elderly assistance. In: 18th proceedings of Intelligent Systems Design and Applications; 2018.
67. Yao Y. An outlie of a theory of three-way decisions. In: Rough sets and current trends in computing (Yao JT, et al., eds,). LNAI 7413 2012, p. 1–17.
68. Veloz T, Razeto P. The state context property formalism: from concept theory to the semantics of music. Soft Comput. 2017;21(6):1505–13.
69. Van Gulick R. Consciousness, The Stanford Encyclopedia of Philosophy, Spring 201: Metaphysics Research Lab, Stanford University; 2018.
70. Yearsley JM. Advanced tools and concepts for quantum cognition: a tutorial. J Math Psychol. 2017;78:24–39.
71. Marvan T, Polák M. Unitary and dual models of phenomenal consciousness. Conscious Cogn. 2017;56(March):1–12.
72. Howard N, Hussain A. The fundamental code unit of the brain: towards a new model for cognitive geometry. Cognit Comput. 2018;10(3):426–36.
73. Dai J, Deng Y. A new method of predict the interference effect in quantum-like Bayesian networks. Soft Comput. 2020;24:10287–94.
74. Vallverdu J, Castro O, Mayene R, Talanov M, Levin M, Baluska F, Gunji Y, Dussutour A, Zenil H, Adamatzky A. Slime mould: the fundamental mechanism of biological cognition. BioSystems. 2018;165:57–70.
75. Mardinoglu A, Nielsen J. Systems medicine and metabolic modelling. J Intern Med. 2012;271:142–54.
76. Friston K. A free energy principle for a particular physics. arXiv. 2019. https://arxiv.org/abs/1906.10184.
77. Parr T. Message passing and metabolism. Entropy. 2021;23:606. https://doi.org/10.3390/e23050606.
78. Basios V, Gunji Y-P. Chaotic dynamics in biological information processing: revisiting and revealing its logic (mini review). Opera Med Physiol. 2017;3(1):1–13.
79. Donoghue T, Haller M, Peterson EJ, Varama P, Sebastian P, Gao R, Noto T, Lala AH, Wallis JD, Knight RT, Shestyuk A, Voytek B. Parameterizing neural power spectra into periodic and aperiodic components. Nat Neurosci. 2020;23:1655–65.
80. , Gunji, Y-P., Kawai, T., Murakami, H., Tomaru, T., Minoura M., Shinohara S. Lévy walk in swarm models based on Bayesian and inverse Bayesian inference. Comput Struct Biotechnol J. 2021;19:247–60.

81. Shinohara S, Manome N, Nakajima Y, Gunji Y-P, Moriyama T, Okamoto H, Mitsuyoshi S, Chung U-I. Power laws derived from a Bayesian decision making model in non-stationary environments. Symmetry. 2021;13(4):718.