

# Interpreting Concept Learning in Cognitive Informatics and Granular Computing

Yiyu Yao *IEEE Member*

**Abstract**—Cognitive informatics and granular computing are two emerging fields of study concerning information and knowledge processing. A central notion to this processing is information and knowledge granularity. Concepts, as the basic units of thought underlying human intelligence and communication, may play a fundamental role when integrating the results from the two fields in terms of information and knowledge coding, representation, communication, and processing. While cognitive informatics focuses on information processing in the abstract, in machines and in the brain, granular computing models such processing at multiple levels of granularity. In this paper, we examine a conceptual framework for concept learning from the viewpoints of cognitive informatics and granular computing. Within the framework, we interpret concept learning based on a layered model of knowledge discovery.

**Index Terms**—Cognitive informatics, granular computing, concept formation, structured thinking, structured problem solving, structured information processing

## I. INTRODUCTION

The study of human intelligence and its applications in problem solving is a topic covered in many disciplines, including philosophy, mathematics and logic, neuroscience, psychology, cognitive science, computer science, and many more. It is perhaps the most controversial area where we have witnessed a wide variety of theories, frameworks, models, and viewpoints. The extensive studies enhance our understanding on the one hand and lead to some confusion on the other. New theories and different interpretations of existing theories of this age-old problem have appeared time and again.

Cognitive informatics [42], [43], [44], [45], [49], [50] and granular computing [1], [2], [11], [14], [25], [28], [29], [32], [52], [53], [55], [56], [59], [60], [62], [63], [65], [71] are two emerging studies on human intelligence and human-inspired problem solving, with different emphases. According to Wang [43], “Cognitive informatics is a transdisciplinary expansion of information science that studies computing and information processing problems by using cognitive science and neuropsychology theories, and studies the cognitive information processing mechanisms of the brain by using computing and informatics theories.” The main focus of cognitive informatics is the study of natural intelligence in general and its information processing mechanisms in particular.

Granular computing is an area of study that explores different levels of granularity in human-centered perception,

problem solving, information processing, as well as their implications and applications in the design and implementation of knowledge-intensive intelligent systems [60], [62], [63], [65]. In some sense, granular computing has a more concrete agenda. It focuses on thinking, problem-solving, and information processing with hierarchical structures showing multiple levels of granularity [65].

An examination of the scopes, goals and methodologies of cognitive informatics and granular computing suggest that there exists a close relationship between both fields [64]. Through a fruitful marriage of the two, it is possible to investigate mechanisms of natural intelligence and information processing as well as their practical applications. To achieve such long-term research goals, it is necessary to have many small and concrete steps. Instead of presenting a full examination of the integration of the two fields, we will interpret concept formation and learning from a unified view of cognitive informatics and granular computing.

Concepts are the basic units of thought that underlie human intelligence and communication. Concept formation and learning is a central topic in both cognitive informatics and granular computing. Wang *et al.* propose a layered reference model of the brain [50], in which concept formation and learning is considered as an important brain function at a conscious process layer. In recent papers, Wang suggests a denotational mathematical representation of concept algebra [48] and gives a description and explanation of concept learning based on the concept algebra [46], [47]. As a complement to these studies, we present a conceptual framework for interpreting concept formation and learning based on a unified view with contributions from cognitive informatics and granular computing. This process gives an example to show that the integration of the two fields is possible at least at a conceptual level.

Concept formation and learning is an important topic of machine learning, knowledge discovery and data mining. In these respective fields of studies, one may find extensive investigations on many approaches, algorithms, and tools. In comparison, less attention is paid to conceptual modeling and explanations. This motivates us to concentrate on a conceptual level examination in the current study.

This paper is an extended version of a conference paper [61]. The latest results from cognitive informatics and granular computing [65] enable us to have explanations of information processing at four levels as follows. Section II covers the first two levels. While cognitive informatics focuses on information processing based on the Object-Attribute-Relation (OAR) model and the concept algebra [47], [48], [51], granular

Yiyu Yao is with the Department of Computer Science, University of Regina, Regina, Saskatchewan, Canada S4S 0A2, E-mail: yyao@cs.uregina.ca. This work is partially supported by a Discovery Grant from NSERC Canada. The author would like to thank Professor Yingxu Wang for his encouragement and anonymous reviewers for their constructive comments when preparing this paper.

computing explores a special type of information structures characterized by multiple levels of granularity. Section III is the third level of explanation with a specific information processing task known as knowledge discovery. This level is further divided into three sub-levels based on a layered framework for knowledge discovery and machine learning, consisting of the philosophy level, the algorithm/technique level, and the application level [58]. Section IV is the most concrete level of explanation, dealing with the issues of concept formation and learning in the three sub-levels of Section III, respectively. As we move to more concrete levels, more detailed results are added.

## II. A UNIFYING VIEW OF COGNITIVE INFORMATICS AND GRANULAR COMPUTING

The main contributing disciplines to cognitive informatics are cognitive science and information science. The cognitive processes of human perception, knowing, and problem solving also serve as a foundation of granular computing. A unifying view of cognitive informatics and granular computing may be obtained by drawing results from cognitive science and information science.

The view consists of two parts, namely, the information processing triangle and the granular computing triangle. The former provides a high-level understanding of the study of information processing from three angles. The latter concretizes the discussion at the next level by considering a special type of information processing with the aid of multiple levels of granularity. In Sections III and IV, we make the discussion to be more specific and concrete at another two levels, one for knowledge discovery in general and the other for concept learning in particular. This level-wise conceptual formulation enables us to see the same problem at four levels of abstraction, with each level focusing on different aspects of the same issues.

### A. The Information Processing Triangle

Cognitive science may be broadly defined as the study of the mind, intelligence and its computational process [36], [40]. It is assumed that human thinking may best be studied based on the notions of representation and process. While representational structures describe the reality, the processes operate on those structures [18], [40], [51]. Simon and Kaplan [36] suggest that cognitive science may be approached in three ways, namely, studying an abstract theory of intelligent processes, studying human or animal intelligence, and studying computer intelligence. This immediately leads to the introduction of the information and knowledge processing triangle, as shown in Fig. 1.

In the triangle, the three corners represent information processing in the abstract (for example, mathematics and logics), in the brain (i.e., natural intelligence of humans and animals), and in machines (i.e., artificial intelligence). The last two topics are in fact reflected by the subtitle of a book by Sowa [39] published in 1984, namely, "information processing in mind and machine." The information processing

triangle will play an important role in the study of cognitive informatics and granular computing.

Studies of the three aspects of information processing are typically the focus of different disciplines. Information processing in the abstract is extensively investigated in philosophy, mathematics and logics. Information processing in machines is studied in computer science and information science in general and artificial intelligence in particular. Information processing in the brain is researched in psychology and neuroscience. By considering the three aspects in a unified framework from the viewpoints of cognitive science, cognitive informatics and granular computing, we have an integrated understanding of information processing.

The three aspects of information processing may be viewed to be both relatively independent and closely related, depending on our standing point. They are independent in the sense that each one of them can be studied without direct reference to the others. They are related in the sense that any one of them supports or implements the other two. Results from any one are complementary to, and may enhance our understanding of, that of the others.

With respect to their relative independence, we may focus on abstract models and architectures without worrying about their implementations. The same abstract theory may be implemented either physically or biologically. In addition, a physical interpretation does not have to mimic the biological implementation, and the former may not be able to explain the latter.

With respect to their close connections, there are bidirectional relationships between any particular aspect and the other two. The processing in the brain and machines may be viewed as biological and physical implementations, respectively, of abstract models and architectures. Abstract models and architectures of information processing may be built based on studies of information processing in the brain and in machines. On the one hand, the fundamental theories, models, and architectures of the computers had been studied before actual computers were built. On the other hand, an in-depth understanding of information processing in machines allows us to build new abstract theories and models of computing. To a large extent, computers have fundamentally changed the ways in which we view and study information processing in the abstract.

The bidirectional relationships between information processing in the abstract and in the brain are not very clear yet. There still does not exist a well-accepted abstract model that describes sufficiently and precisely information processing in the brain. This is a very difficult brain reverse engineering problem [12], which has received much attention recently.

As two different implementations, information processing in the brain and in machines are closely related. Although information processing in machines may simulate or emulate human ways of knowing and thinking at the behavior level, different internal working mechanisms are most likely to be used. In other words, physical and biological implementations may be very different. Nevertheless, we often use one of them to interpret the other. Information processing in machines may help us in understanding processing in the brain, and a better

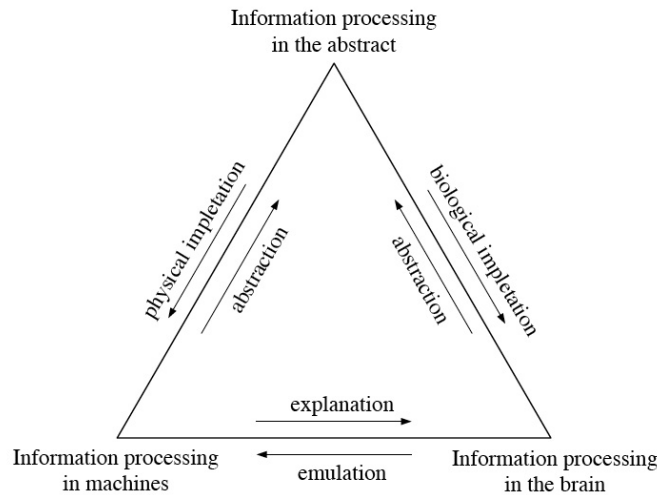


Fig. 1. The information processing triangle

understanding of information processing in the brain may enable us to build better machines for information processing.

There is a growing interest in the study of information processing in the brain in an attempt to overcome limitations of current machine intelligence. It is expected that an understanding of the brain will help us to design and build similar biologically inspired machines [12], [42], [72]. As suggested by Kurzweil [12], an important application is to actually interface brains and computers. In other words, we are seeking for an integrated physical and biological implementation of information processing, in which machines and brains support each other. The findings and an in-depth understanding of human intelligence would have a significant impact on the development of the next generation technologies in informatics, computing, software, and cognitive sciences [12], [42], [72], [73].

### B. Granular Structures

In modeling information processing in the three aspects, granular computing considers special types of representational structures known as granular structures and associated processes [60], [62], [63], [65]. A primitive notion is a granule representing a part of a whole. A real-world problem normally consists of a web of interacting and interrelated parts. Like systems theory [4], granular computing explores the composition of parts and their interrelationships and connections to the whole.

Granular structures can best be explained in terms of granules, levels, and hierarchies [63], [65]. A granule plays two distinctive roles as both a part and a whole. Its particular role is determined by our focal points of observation and investigation. A granule is considered to be a part when we study other granules to which the granule is an element. It is considered to be a whole when we study the family of its constituent granules. This part-whole relationship suggests a partial ordering of granules, which can be described by a hierarchical structure. A granule itself can be, in turn, a

hierarchical structure. We may view a hierarchy as a structure of (partially) ordered multiple levels. Each level is made up of a family of granules of the same type, nature, or details.

A hierarchical structure is built by a vertical separation of levels and a horizontal separation of granules at the same hierarchical level. The separations explore the notion of approximations and a loose coupling of parts [4], [35]. In forming a granule, one may ignore the subtle differences between its elements as well as their individual connections to others. That is, a group of elements may be treated approximately as a whole when studying their relations to others. Each level may be viewed as a representation of a problem at a specific level of granularity. The relationship between levels can be interpreted in terms of abstraction, control, complexity, detail, resolution, etc.

A hierarchy provides a multilevel representation. A problem with a single hierarchy representation is that we only have one particular angle or point-of-view. Although different levels in a hierarchy can also represent different views depending on our cognition, in this paper we refer to them as levels. That is, each view consists of many levels of different granularity. Useful information may be lost with a hierarchy of granules, instead of a web of granules. To resolve this difficulty, we can construct many hierarchies to represent multiple views. Thus, granular structures have the features of multiple levels brought on by a single hierarchy and multiple views brought on by many hierarchies [5], [63].

Many attempts have been made in establishing a model of the brain using hierarchical structures. Hawkins uses the notion of cortical hierarchies for deriving a memory-prediction framework for explaining intelligence [10]. He convincingly argues that the hierarchical organization of the brain is to capture the hierarchical organizations in nature. In his top-down approach for understanding the brain, a model of the cortex is given by highlighting its hierarchical connectivity and the information flows up and down the hierarchy. Wang *et al.* [50] propose a layered reference model of the brain by dividing various brain functions and activities into several

layers. It provides a different hierarchical view of the brain. Kurzweil [12] suggest that we must build models of the brain at the right level. In particular, he considers subneural models, neuromorphic models, and higher-level models. The ideas of levels of description and explanation have also been considered in cognitive science [3].

Although elements in each level and their interpretations are different in these models, the hierarchical structures seem to be a common feature. They provide a cognitive science and cognitive informatics foundation for the adoption of hierarchical structures in granular computing.

### C. The Granular Computing Triangle

The study of granular computing explores the multilevel and multiview granular structures. It can be approached from three perspectives based on the granular computing triangle, as shown in Fig. 2. The triangle integrates the philosophy, the methodology and the computation of a granular method of problem solving and information processing [60], [62], [63], [64], [65]. It promotes a philosophy of structured thinking, a methodology of structured problem solving, and a computation of structured information processing.

From the philosophical perspective, granular computing is a way of structured thinking. It offers structured views for perceiving and understanding the real world using different levels of granularity. Results from two complementary philosophical views about the complexity of real-world problems, i.e., the traditional reductionist thinking and the more recent systems thinking, are drawn together to form the philosophy of granular computing.

According to reductionist thinking, a complex system or problem can be divided into simpler and more fundamental parts, and each part can be further divided. An understanding of the system can be reduced to the understanding of its parts. In contrast, systems thinking shifts from parts to the whole, in terms of connectedness, relationships, and context [4], [13]. A complex system is viewed as a whole consisting of a web of interconnected, interacting, and highly organized parts. The properties of the whole are not present in any of its parts, but emerge from the interactions and relationships of the parts.

The reductionist thinking and systems thinking agree on the modeling of a complex system in terms of the whole and parts. Based on this commonality, granular computing attempts to combine analytical thinking for decomposing a whole into parts and synthetic thinking for integrating parts into a whole. Granular computing stresses the importance of the conscious effects in thinking with the aid of hierarchical structures.

The methodology of granular computing is structured problem solving. It is inspired by human problem solving and can be viewed as an example of natural-inspired computing [16].

Granular computing focuses on systematic approaches, effective principles, and practical heuristics and strategies that have been used effectively by humans for solving real-world problems. A central issue is the exploration of granular structures based on fundamental principles of granular computing. For example, three such principles are [65]:

- The principle of multilevel granularity;

- The principle of focused efforts;
- The principle of granularity conversion.

The first principle emphasizes the importance of modeling in terms of hierarchical structures. Once such structures are obtained, the second principle calls for attention on the focal point at a particular stage of problem solving. The third principle links the different stages in this process. These principles guide the three basic tasks of granular computing, namely, constructing granular structures, working at a particular level of the structure, and moving up and down in the structures.

The computation paradigm of granular computing is the structured information processing paradigm [1]. It focuses on implementing knowledge-intensive systems based on granular structures. Like the study of cognitive science discussed earlier, two related basic notions are representations and processes [18], [40]. Representation covers the formal and precise description of granules and granular structures. Processes may be broadly divided into the two classes: granulation and computation with granules. Granulation involves the construction of granules, levels, and hierarchies. Computation explores the granular structures, involving two-way communications up and down in a hierarchy, as well as switching between levels.

The information processing triangle presents a high-level explanation without constraints on any particular type of information structure. Granular computing enables us to talk about the same problem with an explicit reference to granular structures, namely, a particular class of information structures. The three perspectives of granular computing may be related to the three information processing aspects. It is plausible to link structured thinking to information processing in the abstract, structured problem solving to processing in the brain, and structured information processing to processing in machines. Alternatively, structured thinking may be used to explain information processing in the brain. It may be difficult to separate structured thinking and structured problem solving, as they are in fact inter-weaved tightly together. While structured thinking provides guidelines and leads naturally to structured problem solving, structured problem solving implements the philosophy of structured thinking. The information processing paradigm offered by granular computing will lead to effective and concrete computational models.

As a high-level explanation, each perspective of the granular computing triangle can be further divided into sub-perspectives at a subsequent level of details. This enables us to talk about more specific issues. For example, granular modeling can be discussed at the next level within each of three perspectives, but with different emphasis. Computational models of granular computing and the associated operators can be studied at the next level of detail within the computational perspective. A more complete discussion on this triarchic theory of granular computing can be found in other papers [59], [63], [66], [67].

### III. A LAYERED MODEL FOR KNOWLEDGE DISCOVERY

The unifying view of cognitive informatics and granular computing has significant implications in the understanding, explanation and implementation of intelligent information processing. That is, studying information processing from

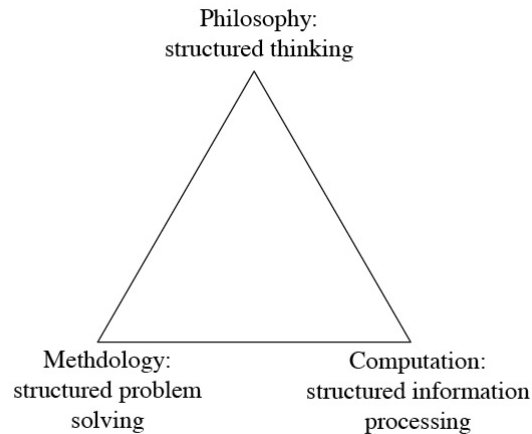


Fig. 2. The granular computing triangle

different aspects and with multiple levels provides a huge hint on the investigation of any particular field. As an illustration, in this section we examine a three-layered framework of knowledge discovery [57], [58], of which concept formation and learning is an example. The three-layered framework is explained based on the ideas emerging from the discussions of previous sections. That is, we seek for multiple levels of understanding of knowledge discovery and data mining. As one of the possible ways to model knowledge discovery and data mining conceptually, this framework enables us to see clearly many fundamental issues.

In the three-layered framework, one separates the study of knowledge and the study of knowledge discovery algorithms, and in turn separates them from the study of the utility of the discovered knowledge. The three levels are made up of the philosophy level, the algorithm/technique level, and the application level [58]. They focus on three fundamental questions. The philosophical level addresses questions about knowledge, the algorithm level concentrates on knowledge discovery methods, and the application level deals with the utility of the discovered knowledge.

The **philosophy level** is the study of knowledge and knowledge discovery methods in the abstract and in the mind. One attempts to answer the fundamental question, namely, what is knowledge? There are many related issues to this question, such as the representation of knowledge in the brain, the expression and communication of knowledge in words and languages, the relationship between knowledge in the brain and in the external world, and the classification and organization of knowledge. One also needs to study the cognitive process in which knowledge is acquired, understood, processed, and used by humans. It is necessary to precisely define “knowledge” and the “basic unit of knowledge,” which serve as the primitive notions of knowledge discovery.

The **algorithm/technique level** is the study of knowledge and knowledge discovery in machines. One attempts to answer the question, how to discover knowledge? There are many issues related to this question, such as the coding, storage,

retrieval of knowledge in a computer, and the implementation of human knowledge discovery methods in programming languages. The focus of the technique level is on algorithms and methods for extracting knowledge from data. The main stream of research in machine learning, data mining, and knowledge discovery has concentrated on the technique level.

The **application level** is about the ultimate goal of knowledge discovery, namely, to effectively use the discovered knowledge. It should focus on the notions of “usefulness” and “meaningfulness” of the discovered knowledge. Although these notions cannot be discussed in total isolation with applications, they may be studied in terms of utility theory and measurement theory [68].

The division between the three levels is not a clear cut, and they may have overlap with each other. It is expected that the results from philosophy level will provide guidelines and set the stage for the algorithm and application levels. On the other hand, it is desirable that philosophical study does not depend on the availability of specific techniques, and technical study is not constrained by a particular application. The existence of a type of knowledge in data is unrelated to whether we have an algorithm to extract it. The existence of an algorithm does not necessarily imply that the discovered knowledge is meaningful and useful.

#### IV. AN INTERPRETATION OF CONCEPT LEARNING

The study of concepts is central to philosophy, psychology, cognitive science, inductive data processing and analysis, and inductive learning [17], [20], [38], [39], [41]. Concepts are assumed to be basic constituents of thought and belief, and the basic units of thought and knowledge. Within the three-layered framework of the last section, we present an interpretation of concept formation and learning.

##### A. The Philosophy Level

There are many views of concepts, concept formation and learning [30], [38], [39], [41]. The classical view treats

concepts as entities with well-defined boundaries and describable by sets of singly necessary and jointly sufficient conditions [41]. Other views include the prototype view, the exemplar view, the frame view, and the theory view [41]. Each view captures specific aspects of concepts, and has a different implication for concept formation and learning. The applications of different views for inductive data analysis have been addressed by many authors [33], [39], [41].

We review the classical view of concepts and the basic issues related to concept formation and organization. More complete treatments of these issues can be found in references [17], [24], [30], [38], [39], [41].

1) *Classical view of concepts*: In the classical view, every concept is understood as a unit of thought that consists of two parts, the intension and the extension of the concept [38], [39], [41]. The intension (comprehension) of a concept consists of all properties or attributes that are valid for all those objects to which the concept applies. The extension of a concept is the set of objects or entities which are instances of the concept. All objects in the extension have the same properties that characterize the concept. In other words, the intension of a concept is an abstract description of common features or properties shared by elements in the extension, and the extension consists of concrete examples of the concept. A concept is thus described jointly by its intension and extension.

Extensional objects are mapped to intensional concepts through perception, and concepts are coded by words in speech. The two mappings of perception and speech define an indirect mapping between words and objects [26], [39]. This is depicted by the meaning triangle, in which one corner represents the concept, intension, thought, idea, or sense, another corner represents symbol or word, and the third corner represent the referent, object, or extension.

The classical view of concepts enables us to study concepts in a logic setting in terms of intensions and also in a set-theoretic setting in terms of extensions. Reasoning about intensions is based on logic [39]. Inductive inference and learning attempt to derive relationships between the intensions of concepts based on the relations between the extensions of concepts. Through the connections between extensions of concepts, one may establish relationships between concepts [57], [58].

2) *Concept formation and structures*: Human knowledge is conceptual and forms an integrated whole. In characterizing human knowledge, one needs to consider two topics, namely, context and hierarchy [30], [37]. The two topics have significant implications for concept formation and organization.

A context in which concepts are formed provides meaningful interpretations of the concepts. The theory view of concepts attempts to, to a large extent, reflect the contextual feature of concepts [24]. It is assumed that the formation of individual concepts and the overall conceptual structure depend on one's theory of a domain. One's theories and complex knowledge structures play a crucial role in concept formation, combination and learning.

Human knowledge is organized in a tower or a partial ordering. The base or minimal elements of the ordering are the most fundamental concepts. The higher-level concepts

depend on lower-level concepts [37]. The first-level concept is formed directly from the perceptual data [30]. The higher-level concepts, representing a relatively advanced state of knowledge, are formed by a process of abstracting from abstractions [30].

In concept formation, there are two basic issues known as aggregation and characterization [9]. Aggregation aims at the identification of a group of objects so that they form the extension of a concept. Characterization attempts to describe the derived set of objects in order to obtain the intension of the concept [9].

For aggregation, one considers two main processes called differentiation and integration [30]. Differentiation enables us to grasp the differences between objects, so that we can separate one or more objects from the others. Integration is the process of putting together elements into an inseparable whole. As the final step in concept formation, characterization provides a definition of a concept.

Humans form and learn concepts based on multiple strategies depending on different interpretations of concepts [31]. Theories for explaining concept learning include the stimulus-response association theory, the hypothesis testing theory, the rule learning theory, and the prototype learning theory.

## B. The Technique Level

Based on the philosophy level study, one can build a computational model for concept formation and learning. A particular concrete model is normally based on some philosophical assumptions and may not be able to cover all issues. As an illustration, we consider a simple model. The intensions are expressed as formulas of a logic language. The extensions are defined by adopting Tarski's approach through the notions of a model and satisfiability [7], [27], [57], [58]. Concept learning is modeled as search in a conjunctive concept space.

1) *Intensions of concepts defined by a language*: Traditionally, the intension of a concept is given by a set of properties. In artificial intelligence, one can define a language so that the intension of a concept is expressed as a formula of the language.

Let  $At$  be a finite set of attributes or features. For each attribute  $a \in At$ , we associate it with a set of values or labels  $V_a$ . Let  $U$  be a universe whose elements are called objects. For each  $a \in At$ , there is a mapping  $I_a$  connecting elements of  $U$  and elements of  $V_a$ . Furthermore, it is assumed that the mapping  $I_a$  is single-valued. In this case, the value of an object  $x \in U$  on an attribute  $a \in At$  is denoted by  $I_a(x)$ . It should be pointed out that we consider only a very restricted language. The main purpose is to illustrate the ideas of the conceptual framework. One may extend the language in real-world applications, based on the same conceptual framework.

In order to formally define intensions of concepts, we adopt the decision logic language  $\mathcal{L}$  used and studied by Pawlak [27]. Formulas of  $\mathcal{L}$  are constructed recursively based on a set of atomic formulas corresponding to some basic concepts. An atomic formula is given by  $a = v$ , where  $a \in At$  and  $v \in V_a$ . For each atomic formula  $a = v$ , an object  $x$  satisfies it if  $I_a(x) = v$ , written  $x \models a = v$ . Otherwise, it does not satisfy

$a = v$  and is written  $\neg x \models a = v$ . From atomic formulas, we can construct other formulas by applying the logic connectives  $\neg$ ,  $\wedge$ ,  $\vee$ ,  $\rightarrow$ , and  $\leftrightarrow$ . The satisfiability of any formula is defined as follows:

- (1)  $x \models \neg\phi$  iff not  $x \models \phi$ ,
- (2)  $x \models \phi \wedge \psi$  iff  $x \models \phi$  and  $x \models \psi$ ,
- (3)  $x \models \phi \vee \psi$  iff  $x \models \phi$  or  $x \models \psi$ ,
- (4)  $x \models \phi \rightarrow \psi$  iff  $x \models \neg\phi \vee \psi$ ,
- (5)  $x \models \phi \leftrightarrow \psi$  iff  $x \models \phi \rightarrow \psi$  and  $x \models \psi \rightarrow \phi$ .

The language  $\mathcal{L}$  can be used to reason about intensions. Each formula represents an intension of a concept. For two formulas  $\phi$  and  $\psi$ , we say that  $\phi$  is more specific than  $\psi$ , and  $\psi$  is more general than  $\phi$ , if and only if  $\models \phi \rightarrow \psi$ , namely,  $\psi$  logically follows from  $\phi$ . In other words, the formula  $\phi \rightarrow \psi$  is satisfied by all objects with respect to any universe  $U$  and any information function  $I_a$ . If  $\phi$  is more specific than  $\psi$ , we write  $\phi \preceq \psi$ , and call  $\phi$  a sub-concept of  $\psi$ , and  $\psi$  a super-concept of  $\phi$ .

2) *Conjunctive concept space*: Concept learning, to a large extent, depends on the structures of concept space and the target concepts. In general, one may not be able to obtain an effective and efficient learning algorithm, if no restrictions are imposed on the concept space. For this reason, each learning algorithm typically focuses on a specific type of concept. For clarity, in this paper we only consider concepts defined by using only the logic connective  $\wedge$ .

Consider the class of conjunctive concepts used in version space learning [21], [22]. Let  $\text{CN}(\mathcal{L})$  denote the class of conjunctive concepts. It contains the special formula  $\top$  which is satisfied by every object, the atomic formula, and formula constructed from atomic formula only by the logic connective  $\wedge$ . Furthermore, we assume that an attribute appears at most once in each formula of  $\text{CN}(\mathcal{L})$ .

The class  $\text{CN}(\mathcal{L})$  is referred to as the conjunctive concept space. For two concepts with  $\phi \preceq \psi$ ,  $\phi$  is called a specialization of  $\psi$  and  $\psi$  a generalization of  $\phi$ . Furthermore,  $\phi$  is called a most general specialization of  $\psi$  and  $\psi$  a most specific generalization of  $\phi$ , if there does not exist another concept between  $\phi$  and  $\psi$ . The conjunctive concept space can be represented as a graph by connecting a concept with its most specific generalizations and its most general specializations. At the top most level, the most general concept is defined by the formula  $\top$ . The next level of concepts are defined by atomic formulas. The combination of two atomic formulas produces the next level of concepts, and so on. Finally, at the bottom most level, a most specific concept is formed by the conjunction of each atomic formula from every attribute.

Figure 3 draws part of the graph of the conjunctive space for three attributes,  $\{\text{Height}, \text{Hair}, \text{Eyes}\}$ , with the following domains:

$$\begin{aligned} V_{\text{Height}} &= \{\text{short}, \text{tall}\}, \\ V_{\text{Hair}} &= \{\text{blond}, \text{dark}, \text{red}\}, \\ V_{\text{Eyes}} &= \{\text{blue}, \text{brown}\}. \end{aligned}$$

In the figure, an atomic formula is simply represented by the

attribute value. For example, the atomic formula  $\text{Height} = \text{short}$  is simply written as *short*.

We can classify conjunctive concepts by the number of atomic concepts in them. A concept involving  $k$  atomic concepts is called a  $k$ -conjunction. Obviously, the most general specializations of  $k$ -conjunction are  $(k+1)$ -conjunctions, and the most specific generalizations of  $k$ -conjunction are  $(k-1)$ -conjunctions.

3) *Extensions of concepts defined by an information table*: In inductive learning and concept formation, extensions of concepts are normally defined with respect to a particular training set of examples. With respect to a dataset, we can build a model based on an information table:

$$M = (U, At, \{V_a \mid a \in At\}, \{I_a \mid a \in At\}). \quad (1)$$

In this case,  $U$  is typically a finite set of objects.

If  $\phi$  is a formula, the set  $m(\phi)$  defined by:

$$m(\phi) = \{x \in U \mid x \models \phi\}, \quad (2)$$

is called the meaning of the formula  $\phi$  in  $M$ . The meaning of a formula  $\phi$  is therefore the set of all objects having the property expressed by the formula  $\phi$ . In other words,  $\phi$  can be viewed as the description of the set of objects  $m(\phi)$ . Thus, a connection between formulas and subsets of  $U$  is established. Obviously, the following properties hold [27]:

- (a)  $m(\neg\phi) = -m(\phi)$ ,
- (b)  $m(\phi \wedge \psi) = m(\phi) \cap m(\psi)$ ,
- (c)  $m(\phi \vee \psi) = m(\phi) \cup m(\psi)$ ,
- (d)  $m(\phi \rightarrow \psi) = -m(\phi) \cup m(\psi)$ ,
- (e)  $m(\phi \equiv \psi) = (m(\phi) \cap m(\psi)) \cup (-m(\phi) \cap -m(\psi))$ ,

where  $-m(\phi)$  denotes the set complement of  $m(\phi)$ . With the introduction of language  $\mathcal{L}$ , we have a formal description of concepts. A subset of  $X \subseteq U$ , representing the extension of a concept, is said to be definable if there is a formula  $\phi$  such that  $m(\phi) = X$ . A concept definable in a model  $M$  is a pair  $(\phi, m(\phi))$ , where  $\phi \in \mathcal{L}$ . More specifically,  $\phi$  is a description of  $m(\phi)$  in  $M$ , the intension of concept  $(\phi, m(\phi))$ , and  $m(\phi)$  is the set of objects satisfying  $\phi$ , the extension of concept  $(\phi, m(\phi))$ .

Table I is an example of an information table, taken from an example from Quinlan [34]. Each object is described by three attributes. The column labeled by Class denotes an expert's classification of the objects.

A concept  $(\phi, m(\phi))$  is said to be a sub-concept of another concept  $(\psi, m(\psi))$ , or  $(\psi, m(\psi))$  a super-concept of  $(\phi, m(\phi))$ , in an information table if  $m(\phi) \subseteq m(\psi)$ . A concept  $(\phi, m(\phi))$  is said to be a smallest non-empty concept in  $M$  if there does not exist another non-empty proper sub-concept of  $(\phi, m(\phi))$ . Two concepts  $(\phi, m(\phi))$  and  $(\psi, m(\psi))$  are disjoint if  $m(\phi) \cap m(\psi) = \emptyset$ . If  $m(\phi) \cap m(\psi) \neq \emptyset$ , we say that the two concepts have a non-empty overlap and hence are related.

A subset of  $X \subseteq U$ , representing the extension of a concept, is said to be conjunctively definable if there is a formula  $\phi \in \text{CN}(\mathcal{L})$  such that  $m(\phi) = X$ . It is important to note that an

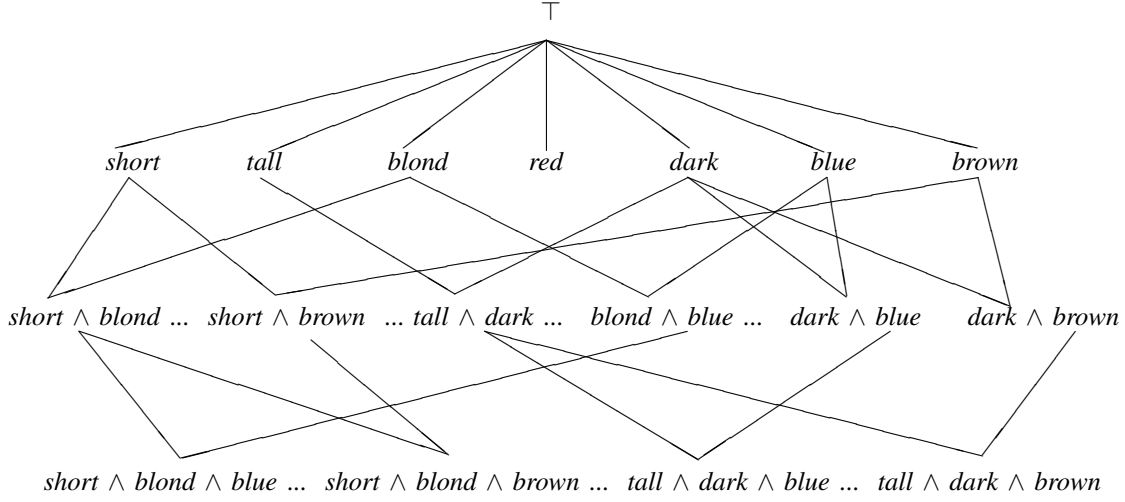


Fig. 3. A conjunctive concept space

Object	Height	Hair	Eyes	Class
$o_1$	short	blond	blue	+
$o_2$	short	blond	brown	-
$o_3$	tall	red	blue	+
$o_4$	tall	dark	blue	-
$o_5$	tall	dark	blue	-
$o_6$	tall	blond	blue	+
$o_7$	tall	dark	brown	-
$o_8$	short	blond	brown	-

TABLE I  
AN INFORMATION TABLE

arbitrary subset of  $U$  may not necessarily be conjunctively definable.

4) *Relationship between concepts in an information table:* Based on the notions introduced so far, we can study a special type of knowledge represented by the relationship between overlapping concepts. This type of knowledge is commonly referred to as a rule. A rule can be expressed in the form,  $\phi \Rightarrow \psi$ , where  $\phi$  and  $\psi$  are intensions of two concepts. A crucial issue is therefore the characterization, classification, and interpretation of rules. It is reasonable to expect that different types of rules represent different kinds of knowledge derivable from a database.

In data mining, rules are typically interpreted in terms of conditional probability [70]. For a rule  $\phi \Rightarrow \psi$ , its characteristics can be summarized by the following contingency table:

	$\psi$	$\neg\psi$	Total
$\phi$	$a$	$b$	$a + b$
$\neg\phi$	$c$	$d$	$c + d$
Total	$a + c$	$b + d$	$a + b + c + d = n$

$$\begin{aligned}
 a &= |m(\phi \wedge \psi)|, & b &= |m(\phi \wedge \neg\psi)|, \\
 c &= |m(\neg\phi \wedge \psi)|, & d &= |m(\neg\phi \wedge \neg\psi)|.
 \end{aligned}$$

Different measures can be defined to reflect various aspects of

rules.

The *generality* of  $\phi$  is defined by:

$$\begin{aligned}
 G(\phi) &= \frac{|m(\phi)|}{|U|} \\
 &= \frac{a + b}{n},
 \end{aligned} \tag{3}$$

which indicates the relative size of the concept  $\phi$ . Obviously, we have  $0 \leq G(\phi) \leq 1$ . A concept is more general if it covers more instances of the universe. A sub-concept has a lower generality than its super-concept. The quantity may be viewed as the probability of a randomly selected element satisfying  $\phi$ .

The *absolute support* of  $\psi$  provided by  $\phi$  is:

$$\begin{aligned}
 AS(\phi \Rightarrow \psi) &= AS(\psi|\phi) \\
 &= \frac{|m(\psi) \cap m(\phi)|}{|m(\phi)|} \\
 &= \frac{a}{a + b},
 \end{aligned} \tag{4}$$

The quantity,  $0 \leq AS(\psi|\phi) \leq 1$ , states the degree to which  $\phi$  supports  $\psi$ . It may be viewed as the conditional probability of a randomly selected element satisfying  $\psi$  given that the element satisfies  $\phi$ . In set-theoretic terms, it is the degree to which  $m(\phi)$  is included in  $m(\psi)$ . Clearly,  $AS(\psi|\phi) = 1$ , if and only if  $m(\phi) \neq \emptyset$  and  $m(\phi) \subseteq m(\psi)$ . That is, a rule with the maximum absolute support 1 is a certain rule.

The *mutual support* of  $\phi$  and  $\psi$  is:

$$\begin{aligned}
 MS(\phi, \psi) &= \frac{|m(\phi) \cap m(\psi)|}{|m(\phi) \cup m(\psi)|} \\
 &= \frac{a}{a + b + c}.
 \end{aligned} \tag{5}$$

One may interpret the mutual support,  $0 \leq MS(\phi, \psi) \leq 1$ , as a measure of the strength of a pair of rules  $\phi \Rightarrow \psi$  and  $\psi \Rightarrow \phi$ .



The *change of support* of  $\psi$  provided by  $\phi$  is defined by:

$$\begin{aligned} CS(\phi \Rightarrow \psi) &= CS(\psi|\phi) \\ &= AS(\psi|\phi) - G(\psi) \\ &= \frac{a}{a+b} - \frac{a+c}{n}. \end{aligned} \quad (6)$$

Unlike the absolute support, the change of support varies from  $-1$  to  $1$ . One may consider  $G(\psi)$  to be the prior probability of  $\psi$  and  $AS(\psi|\phi)$  the posterior probability of  $\psi$  after knowing  $\phi$ . The difference of posterior and prior probabilities represents the change of our confidence regarding whether  $\phi$  is actually related to  $\psi$ . For a positive value, one may say that  $\phi$  is positively related to  $\psi$ ; for a negative value, one may say that  $\phi$  is negatively related to  $\psi$ .

The generality  $G(\psi)$  is related to the satisfiability of  $\psi$  by all objects in the database, and  $AS(\phi \Rightarrow \psi)$  is related to the satisfiability of  $\psi$  in the subset  $m(\phi)$ . A high  $AS(\phi \Rightarrow \psi)$  does not necessarily suggest a strong association between  $\phi$  and  $\psi$ , as a concept  $\psi$  with a large  $G(\psi)$  value tends to have a large  $AS(\phi \Rightarrow \psi)$  value. The change of support  $CS(\phi \Rightarrow \psi)$  may be more accurate.

5) *Concept learning as search*: In concept learning, it is assumed that the extension of a target concept is given through a set of positive and negative examples. However, this subset of objects may not be conjunctively definable. That is, we may not be able to find one formula in  $CN(\mathcal{L})$  to define the target concept. In this case, one may search for the most general concepts in  $CN(\mathcal{L})$  whose extensions are subsets of the extension of the given concept. The conjunctive concept space provides the search space, and the quantitative measures can be used to either direct the search or evaluate the results of learning. Two methods are outlined based on the version space method and Candidate-Elimination algorithms [21], [22], CNF learning [23], PRISM learning method [6], and a granular computing approach for machine learning [54].

We assume that a training set is partitioned into a group of subsets, each represents the extension of a concept. Furthermore, if two objects have the same description, they are in the same class of the partition.

In the first method, we learn all shortest conjunctive formulas that define the sub-concepts of a concept. This can be easily done by searching the conjunctive concept space from general concepts to specific concepts. Figure 4 presents such an algorithm for finding all most general sub-concepts of a family of concepts given by a partition.

For Table I, the algorithm produces the set of conjunctive sub-concepts of the class  $+$ :

$$\begin{aligned} &\{ \text{Hair} = \textit{red}, \\ &\quad \text{Hair} = \textit{blond} \wedge \text{Eyes} = \textit{blue}, \\ &\quad \text{Height} = \textit{short} \wedge \text{Eyes} = \textit{blue}, \\ &\quad \text{Height} = \textit{tall} \wedge \text{Hair} = \textit{blond} \}, \end{aligned}$$

and the set of conjunctive sub-concepts of the class  $-$ :

$$\{ \text{Hair} = \textit{dark}, \quad \text{Eyes} = \textit{brown} \}.$$

From the intension point of view, the two formulas  $\text{Height} = \textit{short} \wedge \text{Eyes} = \textit{blue}$  and  $\text{Hair} = \textit{blond} \wedge \text{Eyes} = \textit{blue}$  do

not have a sub-concept relationship. On the other hand, their extensions with respect to Table I are:

$$\begin{aligned} m(\text{Height} = \textit{short} \wedge \text{Eyes} = \textit{blue}) &= \{o_1\}, \\ m(\text{Hair} = \textit{blond} \wedge \text{Eyes} = \textit{blue}) &= \{o_1, o_6\}. \end{aligned}$$

From the extension point of view, one may choose the second formula as it covers more examples of  $+$ . Therefore, the algorithm in Figure 4 considers both intensions and extensions.

In practice, one may only be interested in a subset of the formulas to characterize a partition. Instead of considering concepts based on the sequence defined by  $k$ , one may consider concepts in a sequence defined based on an evaluation function [54]. The algorithm given in Figure 5 finds a set of the most general concepts whose extensions cover the training sets.

The set of concepts derived from the algorithm in Figure 5 depends on the evaluation function. For example, one may prefer a concept with high generality and high support. The quantitative measures discussed earlier can be used to define various evaluation functions.

Consider the evaluation function defined by:

$$\begin{aligned} eval(\phi) &= \max\{AS(\phi \Rightarrow \text{Class} = +), \\ &\quad AS(\phi \Rightarrow \text{Class} = -)\}. \end{aligned} \quad (7)$$

That is, a concept is evaluated based on its maximum absolute support value of the class  $+$  and the class  $-$ . For the information Table I, the algorithm of Figure 5 produces a set of conjunctive concepts in the following sequence:

$$\begin{aligned} + : &\quad \text{Hair} = \textit{red}, \\ - : &\quad \text{Hair} = \textit{dark}, \\ - : &\quad \text{Eyes} = \textit{brown}, \\ + : &\quad \text{Height} = \textit{short} \wedge \text{Eyes} = \textit{blue}, \\ + : &\quad \text{Height} = \textit{tall} \wedge \text{Hair} = \textit{blond}. \end{aligned}$$

It finds the same atomic concepts as the former algorithm. However, the concepts defined by 2-conjunction are different. When another evaluation function is used, other results may be obtained.

### C. The Application Level

In the application level, one considers the issues related to the correct and effective use of concepts, such as concept definition and characterization, classification, and explanation. One may also explore the relationships between concepts. The application level study of concepts may be guided by the purposes of learning, which in turn can be studied within a wide context of scientific research [69].

Yao and Zhao [69] argued that scientific research and data mining have much in common in terms of their goals, tasks, processes and methodologies. Consequently, data mining and knowledge discovery research can benefit from the long-established studies of scientific research and investigation [19]. Concept learning is a specific topic of data mining and knowledge discovery. The same argument immediately applies.

Scientific research is affected by the perceptions and the purposes of science. Generally speaking, "science is the search

Yao, Y.Y., Interpreting concept learning in cognitive informatics and granular computing  
 IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 39, No. 4, 855-866, 2009.

**Input:** a training set of examples  $S$  and a partition of the training set  $\Pi$ ,  
**Output:** a set of formulas  $F$ .  
**Procedure:** Set  $F = \emptyset$ ;  
 Set  $k = 1$ ;  
 While  $S$  is not empty  
   For each  $k$ -conjunction  $\phi$  which is not a specialization of a concept in  $F$   
     If  $m(\phi) \cap S \neq \emptyset$  is a subset of a class in  $\Pi$   
       Add  $\phi$  to  $F$ ;  
   For each  $k$ -conjunction  $\phi$  in  $F$   
     Delete  $m(\phi)$  from  $S$ ;  
   Set  $k = k + 1$ ;  
 Return ( $F$ ).

Fig. 4. An algorithm for finding all most general concepts for a target partition

**Input:** a training set of examples  $S$ , a partition of the training set  $\Pi$ , and an evaluation function  $eval$ ,  
**Output:** a set of formulas  $F$ .  
**Procedure:** Set  $F = \emptyset$ ;  
 Set  $WF =$  the set of all 1-conjunctions  $\phi$  with  $m(\phi) \neq \emptyset$ ;  
 While  $S$  is not empty  
   Select a best formula  $\phi$  from  $WF$  according to the evaluation function  $eval$   
   If  $m(\phi)$  is a subset of a class in  $\Pi$   
     Add  $\phi$  to  $F$ ;  
     Delete  $\phi$  from  $WF$ ;  
     Delete  $m(\phi)$  from  $S$ ;  
   If  $m(\phi)$  is not a subset of a class in  $\Pi$   
     Replace  $\phi$  by its most general specializations;  
     Delete from  $WF$  concepts that are specializations of concepts in  $F$ ;  
     Delete from  $WF$  every concept  $\phi$  with  $m(\phi) \cap S = \emptyset$ ;  
 Return ( $F$ ).

Fig. 5. An algorithm for finding a set of most general concepts for a target partition

for understanding of the world around us. Science is the attempt to find order and lawful relations in the world. It is a method of viewing the world" [19]. The main purposes of science are to describe and predict, to improve or manipulate the world around us, and to explain our world [19]. The results of the scientific research process provide a description of an event or a phenomenon. The knowledge obtained from research helps us to make predictions about what will happen in the future. Research findings are useful for us to make an improvement in the subject matter or to determine the most effective interventions to bring about desirable changes. Finally, scientists develop models and theories to explain why a phenomenon occurs.

Goals similar to those of scientific research have been discussed by many researchers in data mining. For example, Fayyad *et al.* [8] identified two high-level goals of data mining as prediction and description. Prediction involves the use of some variables to predict the values of some other variables, and description focuses on patterns that describe the data. Some researchers studied the issues of manipulation and action based on the discovered knowledge [15]. A model of explanation-oriented data mining is introduced, which focuses on constructing models for the explanation of data mining results [69]. The ideas may have a significant impact on the

understanding of data mining and effective applications of data mining results.

Concept learning should serve the same purposes, namely, to describe and predict, to improve or manipulate the world around us, and to explain our world.

Consider the example of Table I. Concept learning enables us to describe and explain the classes of + and – using other concepts defined by attributes. Based on the results of the algorithm in Figure 5, the class + is described and explained by the disjunction of three conjunctive concepts:

$$\begin{aligned} &(\text{Hair} = \text{red}) \vee \\ &(\text{Height} = \text{short} \wedge \text{Eyes} = \text{blue}) \vee \\ &(\text{Height} = \text{tall} \wedge \text{Hair} = \text{blond}). \end{aligned}$$

Similarly, the class – is described by the disjunction of two conjunctive concepts:

$$(\text{Hair} = \text{dark}) \vee (\text{Eyes} = \text{brown}).$$

It should be noted that the same class may also be described and explained by a different set of concepts.

For prediction, results of algorithm in Figure 5 produce the

following classification rule:

Hair = *red*  $\Rightarrow$  Class = +,  
 (Height = *short*  $\wedge$  Eyes = *blue*)  $\Rightarrow$  Class = +,  
 (Height = *tall*  $\wedge$  Hair = *blond*)  $\Rightarrow$  Class = +,  
 Hair = *dark*  $\Rightarrow$  Class = - ,  
 Eyes = *brown*  $\Rightarrow$  Class = - .

That is, we can predict the class of an object based on its attribute values.

In some situations, the tasks of description and prediction may not be clearly separated. In order to have a good prediction, one must have a good description and explanation.

The concept learning methods can be applied to study relationships between attributes. This can be simply done by generating the partition  $\Pi$  using one subset of attributes, and by learning using another subset of attributes. The results can be explained in a similar manner.

## V. CONCLUSION

The two emerging fields of study, namely, cognitive informatics and granular computing, are closely related. While cognitive informatics investigates natural intelligence in general, granular computing studies human problem solving based on granular structures. Since both of them concern natural intelligence and its applications, their combination will produce insights and useful results.

A unifying view of cognitive informatics and granular computing consists of the information processing triangle and the granular computing triangle. The former concerns information processing in the abstract, in the mind, and in machines at a higher level without an explicit reference to the information structures used. The latter represents the study of the philosophy, the methodology and the computation of granular computing at a more concrete level, focusing on hierarchical granular structures.

Based on the unified view, we further explore two additional concrete levels. A three-layered framework of knowledge discovery, consisting of the philosophy level, the technique level, and the application level. An interpretation of concept learning is given within this framework. In the philosophy level study, we focus on the definition, interpretation of concepts, and cognitive process for concept formation and learning. In the technique level, we focus on a specific language for defining concepts and present two algorithms for concept learning. In the application level, we study explanations and uses of the learned results.

The objective of the paper is aimed at a more general framework for concept formation and learning from the viewpoints of cognitive informatics and granular computing, rather than a more efficient algorithm. The re-interpretation of existing results leads to a four-level explanation of information processing, namely, from general information processing to granular computing, to knowledge discovery, and to concept formation and learning. As future research, we will apply the proposed framework to study other topics in cognitive informatics.

## REFERENCES

- [1] Bargiela, A. and Pedrycz W. *Granular Computing: An Introduction*, Kluwer Academic Publishers, Boston, 2002.
- [2] Bargiela, A. and Pedrycz, W. The roots of granular computing, *Proceedings of 2006 IEEE International Conference on Granular Computing*, 806-809, 2006.
- [3] Bechtel, W. Levels of description and explanation in cognitive science, *Minds and Machines*, **4**, 1-25, 1994.
- [4] Capra, F. *The Web of Life*, Anchor Books, New York, 1997.
- [5] Chen, Y.H. and Yao, Y.Y. A multiview approach for intelligent data analysis based on data operators, *International Journal of Information Sciences*, **178**, 1-20, 2008.
- [6] Cendrowska, J. PRISM: an algorithm for inducing modular rules, *International Journal of Man-Machine Studies*, **27**, 349-370, 1987.
- [7] Demri, S. and Orłowska, E. Logical analysis of indiscernibility, in: Orłowska, E. (Ed.), *Incomplete Information: Rough Set Analysis*, Physica-Verlag, Heidelberg, 347-380, 1998.
- [8] Fayyad, U.M., Piatetsky-Shapiro, G. and Smyth, P. From data mining to knowledge discovery: an overview, in: Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R. (Eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI/MIT Press, Menlo Park, CA, 1-34, 1996.
- [9] Feger, H. and de Boeck, P. Categories and concepts: introduction to data analysis, in: van Mechelen, I., Hampton, J., Michalski, R.S. and Theuns, P. (Eds.), *Categories and Concepts, Theoretical Views and Inductive Data Analysis*, Academic Press, New York, 203-223, 1993.
- [10] Hawkins, J. (with Blakeslee, S.) *On Intelligence*, Henry Holt and Company, New York, 2004.
- [11] Inuiguchi, M., Hirano, S. and Tsumoto, S. (Eds.) *Rough Set Theory and Granular Computing*, Springer, Berlin, 2003.
- [12] Kurzweil, R. *The Singularity is Near*, Penguin Books, New York, 2005.
- [13] Laszlo, E. *The Systems View of the World: The Natural Philosophy of the New Developments in the Science*, George Brasiller, New York, 1972.
- [14] Lin, T.Y., Yao, Y.Y. and Zadeh, L.A. (Eds.) *Data Mining, Rough Sets and Granular Computing*, Physica-Verlag, Heidelberg, 2002.
- [15] Ling, C.X., Chen, T., Yang, Q. and Cheng, J. Mining optimal actions for profitable CRM, *Proceedings of 2002 IEEE International Conference on Data Mining*, 767-770, 2002.
- [16] Liu, J. and Tsui, K.C. Toward nature-inspired computing, *Communications of the ACM*, **49**, 59-64, 2006.
- [17] Margolis, E. and Laurence, S. Concept, *Stanford Encyclopedia of Philosophy*, 2006, <http://plato.stanford.edu/entries/concepts/> (accessed November 24, 2007).
- [18] Marr, D. *Vision, A Computational Investigation into Human Representation and Processing of Visual Information*, W.H. Freeman and Company, San Francisco, 1982.
- [19] Martella, R.C., Nelson, R. and Marchand-Martella, N.E. *Research Methods: Learning to Become a Critical Research Consumer*, Allyn and Bacon, 1999.
- [20] Michalski, R.S., Carbonell, J.G., and Mitchell, T.M. (Eds.), *Machine Learning, an Artificial Intelligence Approach*, Morgan Kaufmann Publishers, Inc., Los Altos, California, 1983.
- [21] Mitchell, T.M. *Machine Learning*, McGraw-Hill, New York, 1997.
- [22] Mitchell, T.M. Generalization as search, *Artificial Intelligence*, **18**, 203-226, 1982.
- [23] Mooney, R.J. Encouraging experimental results on learning CNF, *Machine Learning*, **19**, 79-92, 1995.
- [24] Murphy, G.L. Theories and concept formation, in: van Mechelen, I., Hampton, J., Michalski, R.S. and Theuns, P. (Eds.), *Categories and Concepts, Theoretical Views and Inductive Data Analysis*, Academic Press, New York, 173-200, 1993.
- [25] Nguyen, H.S., Skowron, A., and Stepaniuk, J. Granular computing: a rough set approach, *Computational Intelligence*, **17**, 514-544, 2001.
- [26] Ogden, C.K. and Richards I.A. *The Meaning of Meaning: a Study of the Influence of Language upon Thought and of the Science of Symbolism*, 8th edition, Harcourt Brace, New York, 1946.
- [27] Pawlak, Z. *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Boston, 1991.
- [28] Pawlak, Z. Granularity of knowledge, indiscernibility and rough sets, *Proceedings of 1998 IEEE International Conference on Fuzzy Systems*, 106-110, 1998.
- [29] Pedrycz, W. (Ed.) *Granular Computing: An Emerging Paradigm*, Physica-Verlag, Heidelberg, 2001.
- [30] Peikoff, L. *Objectivism: the Philosophy of Ayn Rand*, Dutton, New York, 1991.
- [31] Pettijohn, T.F. *Psychology: A ConnecText*, 4th Edition, McGraw-Hill, Columbus, 1998.

Yao, Y.Y., Interpreting concept learning in cognitive informatics and granular computing

IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 39, No. 4, 855-866, 2009.

- [32] Polkowski, L. A model of granular computing with applications: granules from rough inclusions in information systems, *Proceedings of 2006 IEEE International Conference on Granular Computing*, 9-16, 2006.
- [33] Posner, M.I. (Ed.) *Foundations of Cognitive Science*, The MIT Press, Cambridge, Massachusetts, 1989.
- [34] Quinlan, J.R. Learning efficient classification procedures and their application to chess endgames, in: Michalski, J.S., Carbonell, J.G., and Mitchell, T.M. (Eds.), *Machine Learning: An Artificial Intelligence Approach*, vol. 1, Morgan Kaufmann, Palo Alto, CA, 463-482, 1983.
- [35] Simon, H.A. *The Sciences of the Artificial*, The MIT Press, Massachusetts, 1990.
- [36] Simon, H.A. and Kaplan, C.A. Foundations of cognitive science, in: Posner, M.I. (Ed.) *Foundations of Cognitive Science*, The MIT Press, Cambridge, Massachusetts, 1-47, 1989.
- [37] Simpson, S.G. What is foundations of mathematics? 1996, <http://www.math.psu.edu/simpson/hierarchy.html>, retrieved November 21, 2003.
- [38] Smith, E.E. Concepts and induction, in: Posner, M.I. (ed.), *Foundations of Cognitive Science*, The MIT Press, Cambridge, Massachusetts, 501-526, 1989.
- [39] Sowa, J.F. *Conceptual Structures, Information Processing in Mind and Machine*, Addison-Wesley, Reading, Massachusetts, 1984.
- [40] Thagard, P. Cognitive science, *Stanford Encyclopedia of Philosophy*, 2007, <http://plato.stanford.edu/entries/cognitive-science/> (accessed November 24, 2007).
- [41] van Mechelen, I., Hampton, J., Michalski, R.S. and Theuns, P. (Eds.), *Categories and Concepts, Theoretical Views and Inductive Data Analysis*, Academic Press, New York, 1993.
- [42] Wang, Y. Cognitive informatics: a new transdisciplinary research field, *Brain and Mind: A Transdisciplinary Journal of Neuroscience and Neurophilosophy*, **4**, 115-127, 2003.
- [43] Wang, Y. On cognitive informatics, *Brain and Mind: A Transdisciplinary Journal of Neuroscience and Neurophilosophy*, **4**, 151167, 2003.
- [44] Wang, Y. *Software Engineering Foundations, A Software Science Perspective*, Auerbach Publications, New York, 2007.
- [45] Wang, Y. The theoretical framework of cognitive informatics, *The International Journal of Cognitive Informatics and Natural Intelligence*, **1**, 1-27, 2007.
- [46] Wang, Y. The theoretical framework and cognitive process of learning, *Proceedings of the 6th IEEE International Conference on Cognitive Informatics*, 470-479, 2007.
- [47] Wang, Y. The OAR model of neural informatics for internal knowledge representation in the brain, *International Journal of Cognitive Informatics and Natural Intelligence*, **1**, 68-82, 2007.
- [48] Wang, Y. On concept algebra, a denotational mathematical structure for knowledge and software modeling, *The International Journal of Cognitive Informatics and Natural Intelligence*, **2**, 1-19, 2008.
- [49] Wang, Y., Johnston, R.H. and Smith, M.R. (Eds.) (2002) *Proceedings of the 1st IEEE International Conference on Cognitive Informatics*, IEEE Computer Society Press, Los Alamitos, 2002.
- [50] Wang, Y., Wang, Y., Patel, S. and Patel, D. A layered reference model of the brain (LRMB), *IEEE Transactions on Systems, Man, and Cybernetics (C)*, **36**, 124-133, 2006.
- [51] Wang, Y., Zadeh, L.A. and Yao, Y.Y. On the system algebra foundations for granular computing, *International Journal of Software Science and Computational Intelligence*, **1**, 64-86, 2009.
- [52] Yao, J.T. Information granulation and granular relationships, *Proceedings of the 2005 IEEE International Conference on Granular Computing*, 326-329, 2005.
- [53] Yao, J.T. A ten-year review of granular computing, *Proceedings of the 2007 IEEE International Conference on Granular Computing*, 734-739, 2007.
- [54] Yao, J.T. and Yao, Y.Y. A granular computing approach to machine learning, *Proceedings of the 1st International Conference on Fuzzy Systems and Knowledge Discovery*, 732-736, 2002.
- [55] Yao, Y.Y. Granular computing: basic issues and possible solutions, *Proceedings of the 5th Joint Conference on Information Sciences*, 186-189, 2000.
- [56] Yao, Y.Y. Information granulation and rough set approximation, *International Journal of Intelligent Systems*, **16**, 87-104, 2001.
- [57] Yao, Y.Y. Modeling data mining with granular computing, *Proceedings of the 25th Annual International Computer Software and Applications Conference*, 638-643, 2001.
- [58] Yao, Y.Y. A step towards the foundations of data mining, in: Dasarthy, B.V. (Ed.), *Data Mining and Knowledge Discovery: Theory, Tools, and Technology V*, The International Society for Optical Engineering, 254-263, 2003.
- [59] Yao, Y.Y. A partition model of granular computing, *LNCS Transactions on Rough Sets I*, LNCS 3100, 232-253, 2004.
- [60] Yao, Y.Y. Granular computing, *Computer Science (Ji Suan Ji Ke Xue)*, **31**, 1-5, 2004.
- [61] Yao, Y.Y. Concept formation and learning: a cognitive informatics perspective, *Proceedings of the 3rd IEEE International Conference on Cognitive Informatics*, 42-51, 2004.
- [62] Yao, Y.Y. Perspectives of granular computing, *Proceedings of the 2005 IEEE International Conference on Granular Computing*, Vol. 1, 85-90, 2005.
- [63] Yao, Y.Y. Three perspectives of granular computing, *Journal of Nan-chang Institute of Technology*, **25**, 16-21, 2006.
- [64] Yao, Y.Y. Granular computing and cognitive informatics, *Proceedings of the 5th IEEE International Conference on Cognitive Informatics*, 17-18, 2006.
- [65] Yao, Y.Y. The art of granular computing, *Proceeding of the International Conference on Rough Sets and Emerging Intelligent Systems Paradigms*, LNAI 4585, 101-112, 2007.
- [66] Yao, Y.Y. A unified framework of granular computing, in: Pedrycz, W., Skowron, A. and Kreinovich, V. (Eds.), *Handbook of Granular Computing*, Wiley, 401-410, 2008.
- [67] Yao, Y.Y. Granular computing: past, present and future, *Proceedings of the 2008 IEEE International Conference on Granular Computing*, 80-85, 2008.
- [68] Yao, Y.Y., Chen, Y.H. and Yang, X.D. A Measurement-theoretic foundation for rule interestingness evaluation, in: Lin, T.Y.; Ohsuga, S.; Liau, C.-J.; Hu, X. (Eds.) *Foundations and Novel Approaches in Data Mining*, Springer, Berlin, 41-59, 2005.
- [69] Yao, Y.Y. and Zhao, Y. Explanation-oriented data mining, in: Wang, J. (Ed.), *Encyclopedia of Data Warehousing and Mining*, Idea Group Inc., 492-497, 2005.
- [70] Yao, Y.Y. and Zhong, N. An analysis of quantitative measures associated with rules, *Proceedings of PAKDD'99*, 479-488, 1999.
- [71] Zadeh, L.A. Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems*, **19**, 111-127, 1997.
- [72] Zhong, N. Impending brain informatics research from web intelligence perspective, *International Journal of Information Technology and Decision Making*, **5**, 713-728, 2006.
- [73] Zhong, N. Ways to develop human-level web intelligence: a brain informatics perspective, *The Semantic Web: Research and Applications, 4th European Semantic Web Conference*, LNAI 4519, 27-36, 2007.