

Factors Space and its Relationship with Formal Conceptual Analysis: A General View

H. Liu, I. Dzitac, S. Guo

Haitao Liu*

1. Institute of Intelligence Engineering and Mathematics
Liaoning Technical University, Fuxin 123000, China
2. College of Science
Liaoning Technical University, Fuxin 123000, China
*Corresponding author: liuhaitao@lntu.edu.cn

Ioan Dzitac

1. Aurel Vlaicu University of Arad
310330 Arad, Elena Dragoi, 2, Romania
ioan.dzitac@uav.ro
2. Agora University of Oradea
410526 Oradea, P-ta Tineretului 8, Romania,
idzitac@univagora.ro

Sicong Guo

1. Institute of Intelligence Engineering and Mathematics
Liaoning Technical University, Fuxin 123000, China
2. College of Science
Liaoning Technical University, Fuxin 123000, China
guosizong@tom.com

Abstract: Conceptual generation is a key point and basic problem in artificial intelligence, which has been probed in the Formal Concept Analysis (FCA) established by G. Wille. Factors Space (FS) is also a branch of cognition math initiated by P.Z. Wang at the end of last century, which has been applied in information processing with fuzzy concepts effectively. This paper briefly introduces the historic background of FS and its relationship with FCA. FS can be seen as a good partner of FCA on conceptual description and structure extraction; combining FCA with FS, we can get more clear and simple statements and more fast algorithms on conceptual generation.

Keywords: Factors Space (FS), Formal Concept Analysis (FCA), rough sets, conceptual generation, basic concept semi-lattice, fuzzy logic.

1 Introduction

There were three mathematical branches describing cognition and thought of human beings simultaneously: Formal Concept Analysis [32] (Wille 1982), Rough Sets [14] (Pawlak, 1982) and Factors space [29] (Wang, 1982). We call them the branches of Cognition Math since there may be not any mathematical branch clearly declaring its object is to describe cognition process before. Even though many mathematical branches, such as mathematical logic, probability, optimization have been applied widely in artificial intelligence, those branches just do AI jobs spontaneously by their own natural characteristics, while cognition math conscientiously does AI jobs, fitting the needs of the era.

Relying on the involution between intension and extension, R. Wille gave a mathematical definition to the concept and gave algorithms for the generation of conceptual lattice, and since then, the computer has done conceptual generation automatically. He had opened the door of

cognition math. Similarly, Z. Pawlak brought in the rough sets to do knowledge discovering for databases, and P. Z. Wang brought in the factors space to represent things and thinking.

The articles of Wille are very serious. Unfortunately, he focuses on the attribute value rather than the attribute name. The number of columns in his formal background table exponentially increases, and the algorithms on concept lattice in FCA encounters the trap of N-hard complexity. Z. Pawlak emphasizes attribute name rather than attribute value, making the number of columns of formal background table greatly reduced and turned into the table of information system in rough sets, which plays the theoretical foundation for relational database. Unfortunately, his theory is not deep enough to avoid N-hard trap in attribute induction. Pawlak School emphasizes attribute name, but is not really aware of the significance of the attribute name. What is the difference between attribute value and attribute name? Attribute values, such as tall, middle and short, are a group of varying states under the name Height. From the viewpoint of mathematics, an attribute value is a constant, while the name is a variable that represents each one of its states. Height is an important factor for basketball player selection since the selection limits the varying of height. For the same reason, it is important for weightlifter selection, while it is not important for student examination since there is no limit to its varying. To manifest the influence of height with respect to the talents selection, we would like to say that height is a factor. In this paper, the word factor stands for a name, which influences or causes something.

People engaged in database know FCA and RS very well, but they are not familiar with factors space when FS was presented by P. Z. Wang at 1982, which was used to do ontology and cognition rather than database. To make mathematical discussion on randomness vs fuzziness, Wang presented FS to represent the basic space in a probability field and represent the discussion universe in fuzzy field respectively, and discovered a dual relationship: A fuzzy set defined on universe U can be represented by a probability field in the power of U ; fuzzy measures can be determined by the falling shadow of a random set. This is the core of falling shadow theory [18]. Wang has proved the existence and uniqueness theorem around the probability distribution in the power $P(U)$, which lays the foundation for the applications of fuzzy sets and other subjective non-additive measures. FS plays a very important role in the development of fuzzy mathematics in China. Since 2012, Wang has turned his target to data science and finds that by means of the supports of FS, we can unify and deepen the theories of FCA and RS. All subjects in FCA and RS can be briefly and clearly stated, and the mentioned N-hard traps can be resolved by faster algorithms. FS provides natural base for cognition, information and data science. In this paper, we will introduce and develop the theory of FS briefly. We will introduce the historic background of factors space theory in Section 2. What is factors and factors space? We will answer the questions in Section 3. The core idea of background relation of FS will be given in Section 4. In section 5, we will deepen the conceptual generation and its AI applications based on FS. Section 6 will draw a brief conclusion.

2 The historic background of factors space theory

In 1982, Wang presented the first paper of factors space to explore and compare two kinds of uncertainty: fuzziness and randomness. He used factors space to represent basic space in probability and the discussion universe in fuzzy sets respectively. He clearly answered questions about the sources, and distinguished and relationship of the two kinds of uncertainties. We will introduce the historic background of factors space in following sub-sections.

2.1 Basic space in probability

What is randomness? Why and when does it occur? Tossing a coin, we can get two outcomes, Head or Back randomly. To know why randomness occurs, we need to find out those factors that influence the tossing process, such as Shape of coin, Actions of fingers, Condition of desktop, Environmental influence and so on. They will impact the outcomes. We may make such a determinism hypothesis: When the states of all factors are fixed, the outcome will be determinate. If not, there must be some influential factors that have not been considered, and then the hypothesis should be held considering all missing factors. Each factor f has a state space $X(f)$ as a dimension; let F_o be the set of all factors, and they form the high-dimensional space $X(F_o)$, which is what we call the factors space. This hypothesis tells us that, if the considered set F_o of factors is sufficient and if the state of factors can be observed and controlled precisely, then there is no randomness; randomness occurs if the considered factors are not complete or even if complete, they cannot be observed and controlled precisely. Randomness is the uncertainty on the occurrence of event caused by the lack of conditional factors with their observation and controlling.

If we cannot observe a state \mathbf{x} precisely into a single point but a subset C in $X = X(F_o)$, C is called the condition of experiment. Under the condition, we only know \mathbf{x} is in C but don't know where it is. When C is not small enough and crosses the border of the opposite sides with respect to an event, the randomness occurs. Even though conditional factors are not sufficient, there exists causality rule between condition and result. Probability is generalized causality, which is the essential scale of the frequency of an event. The basic model of probability statistics is: Fix the circle, move the point, where circle stands for the condition C , and point stands for the point \mathbf{x} in X .

Kolmogorov gave an axiomatic definition on probability field (Ω, \mathcal{A}, p) , where Ω is the basic space, \mathcal{A} is an σ -algebra on Ω , and the probability p is an additive measure defined on \mathcal{A} . Wang admires him so much that he defines the random variable as a mapping $\xi : \Omega \rightarrow R$; based on the definition, classical probabilities are carried by ξ and become probability distribution functions or distribution densities on real line R , and then probability becomes a modern mathematical branch. However, how can we define a random variable ξ , a mapping from Ω to R ? Only factors space can describe a random variable by a determinate mapping, and the basic space Ω must be a factors space [17]. As a great mathematician, Kolmogorov has had the idea of factors space already.

Viewing the basic space Ω as a factors space, we can study probability from a new aspect: Focusing on the transformation between randomness and certainty. The core problem is how to decrease the randomness so that it is transferred into certainty?

We can divide the set F_o of factors definition into three subset $F_o = G_1 + G_2 + G^c$, where $G_1 + G_2$ is the set of considered factors, and G^c is the set of unconsidered factors, G_1 is the set of considered factors whose states can be observed and controlled with accuracy C , which is small enough so that it does not cross the borders of opposite side of an events. G_2 is the factors in G but not in G_1 . Then we can have a formula

$$\xi = g_1(\mathbf{x}) + g_2(\mathbf{x}) + g^c \tag{1}$$

where g_1 is a function, g_2 is a probability density under the condition C , and g^c is a Gaussian distribution of white noise when all unconsidered factors are neglect-able.

Let ξ be the hitting score of a shooter. We need to calculate its mean $a = M\xi$ and the mean square error $\sigma = (\xi - M\xi)^{1/2}$. Passing through special trains, we can reduce the error σ , and then move the parameter a to be the real target. An important task of probability is how

to promote the hitting probability of a shooter throughout special trains on improving the two parameters a and σ respectively.

2.2 Discussion universe of a fuzzy set

L. A. Zadeh (1921-2017) was a great world-renowned computer scientist in the history [4]. To open a path to develop artificial intelligence from the mathematical base, he presented fuzzy sets in 1965 [36], deeply influencing the progress of AI applications in the world. A fuzzy subset of the universe U is defined by a mapping $\mu_A : U \rightarrow [0, 1]$, called membership function of A . He encouraged Wang to make a study on the discussion universe U ; while the selection of U is crucial for the fuzzy description. For example, "young" is a fuzzy concept. To see whether a man is young, it is very difficult to measure the membership under the factor Age, but is much easier if we consider more factors such as "face", "action", "vigor". P.Z. Wang treated the discussion universe U as a factors space and found that fuzziness is the uncertainty on the conceptual extension caused by the lack of cognitive factors [18]. Even though cognitive factors are not sufficient, there exists the law of excluded middle in recognition still. Membership degree reflects the generalized law of excluded middle, which essentially scales the frequency of covering.

Lacked factors are essentially the objective factors in randomness but essentially the subjective factors in fuzziness. The basic model of probability statistics is: Fix the circle, move the point; the basic model in fuzzy statistics is: Fix the point, move the circle. Where the point stands for a state in factors space X , the circle stands for a subset in X . This is a duality. A circle of U is a point in the power $P(U)$; while a point \mathbf{x} in U can be transferred a circle $\mathbf{x}^* = \{C | C \ni \mathbf{x}\}$ in $P(U)$. Therefore, we get a dual principle between fuzziness and randomness: A fuzzy experience model in the ground U can be transferred to a probability experience model in the sky $P(U)$. According to the principle of duality, Wang presented the falling shadow theory, which defines the membership degree of A as the covering probability of the related random set ξ :

$$\mu_A(\mathbf{x}) = p(\xi \ni x) \quad (2)$$

Falling shadow theory is the population theory for fuzzy statistics (interval statistics, set-valued statistics), which is a special contribution of factors space for fuzzy sets theory praised by the founder L. A. Zadeh [36].

Factors space theory gave subjective measures, including fuzzy measure, a united statement and gave a deep theorem: For any subset A of U , set $A^* = \{C | C \supseteq A\}$, $*A = \{C | C \subseteq A\}$ and called the idea and filter of A respectively. Set $(A^*)^c = \{C | C^c \cap A \neq \emptyset\}$, $(*A)^c = \{C | C \cap A^c \neq \emptyset\}$. Let (X, T) be a topological space, denote $T_1 = \{A^* | A \in T\}$, $T_2 = \{*A | A \in T\}$, $T_3 = \{(A^*)^c | A \in T\}$, $T_4 = \{(*A)^c | A \in T\}$, and let $\mathbf{T}_i (i = 1, 2, 3, 4)$ be the hyper-topologies generated by T_1 , T_2 , T_3 and T_4 respectively. $(P(T), \mathbf{T}_i) (i = 1, 2, 3, 4)$ are hyper-topological spaces on the power of $i = 1, 2, 3, 4$, let \mathbf{A}_i be the σ -algebra generated from \mathbf{T}_i respectively, $(P(T), \mathbf{A}_i) (i = 1, 2, 3, 4)$ are called four basic hyper-measurable spaces on the power $P(T)$ respectively. Let \mathbf{p} be a probability defined on $(P(T), \mathbf{A}_i)$, for any $A \in T$ set

$$b(A) = \mathbf{p}(A^*), n(A) = \mathbf{p}(*A), an(A) = \mathbf{p}((A^*)^c), ab(A) = \mathbf{p}((*A)^c) \quad (3)$$

where b , n , an and ab are known as the belief, plausibility, anti-plausibility and anti-belief measures respectively. They are non-additive measures.

Theorem 1. (Extension and Uniqueness Theorem) *Given any one of the four non-additive measure on A , there exists one and only one probability \mathbf{p} on $(P(T), \mathbf{A}_i)$ such that Eq.(2) is held.*

Proof is given in [18].

Why did Wang use factors space to study on randomness vs randomness? He wants to develop cognition math. Fuzzy sets theory brings fuzzy concepts into math so that the programs can do smart recognition and control like brain of human being; fuzzy sets and fuzzy logic open a new door of mathematics toward AI applications. Factors space can deepen the researches of fuzzy sets. Treating the discussion universe U as a factors space, we can select factors to make a concept clearer and decrease its fuzziness, and we can calculate membership degree more reasonably.

Fuzzy sets represent cognition and thinking from the aspect of extension only, but lack the aspect of intension. Factors space is a plate to combine extension and intension, which can represent things and thinking. H. X Li and others use factors space to do fuzzy computer [25] and fuzzy controller with 4 stages inverse pendulum and factor space has been successfully applied to AI in China.

3 Factors and factors spaces

Based on the ideological background from history, we introduce and develop the factors theory. A lot of theoretical papers before 2012 can be found in the reference [3, 6–11, 13, 16, 19, 20, 24, 26, 31, 34, 35]. Since 2012, FS has turned to the data science [1, 2, 12, 21–23, 27, 28, 30, 37, 38].

We have mentioned that a factor causes something. But according to the interpretation of dictionary, factor is the reason causing something or the essence of thing, which tells us that a factor essences the very thing. What is the thing? Human being has accumulated a lot of knowledge on the generation of things, and the perfect model occurs in biology. Gene is the key to opening the door of biology, which induces the discovery of DNA, the mystery of life. Gene was called the factor by G. Mendel; factor is the generalization of gene, which is the key to opening the door of ontology and cognition, said Wang, who gives special emphasize from philosophy: Anything is united in the opposite of quality and quantity; factor is the root of quality, as gene is the root of biological attributes. Without the string of factors, attributes are like broken strings of pearls sprinkled over a floor [28]. By means of factors, things are organized and concepts are generated.

The group of states bunched by factor f forms a dimension $X(f)$ called the state space of f . Any factor must be defined on a discussion universe U . A factor f can be mathematically defined as a mapping $f : U \rightarrow X(f)$.

Factor is the aspect we observe and focus on; factors can be operated from rough to fine or from fine to rough, reflecting the analysis process and the synthesis process in thinking respectively. Consider a group of simple factors f_1, \dots, f_n defined on U , and they combine to a complex factor denoted as $F_o = \{f_1, \dots, f_n\} = f_1 \vee \dots \vee f_n$. The state space of F_o can be defined as the Cartesian product space $X(F_o) = \prod_i X(f_i)$. Each simple factor f_i performs analysis, which maps an object u in U to $f_i(u)$ in $X(f_i)$; the complex factor F_o performs synthesis and get the coordinate of u in $X(F_o)$. For example, consider Height(f_1), Weight(f_2), Age(f_3) and Sex(f_4) defined on a group of people U . Suppose that $u=John$ in U , $f_1(John)=1.75m$, $f_2(John)=75kg$, $f_3(John)=25years$, $f_4(John)=Male$, then, we get that $F_o(John)=(1.75m, 75kg, 25years, Male)$. John can be represented as a point in the coordinate system. Factors space is a generalized coordinate system of things.

Definition 2. Let $F = (F; \vee \wedge, \mathbf{0})$ be a lattice with minimum $\mathbf{0}$, $X_F = \{X(f)\}_{(f \in F)}$ is a family of sets satisfying

- (1) $X(\mathbf{0}) = \{\emptyset\}$;
- (2) For any irreducible subset T of F ($s, t \in T$ implies $s \wedge t = \mathbf{0}$), we have that

$$X(T) = \prod_{f \in T} X(f) \quad (4)$$

where \prod stands for Cartesian product. $\psi = (U, X_F)$ is called a factors space on U , if for any $u \in U$ and $f \in F$, there exists a unique $f(u) \in X(f)$. (i.e., each $f \in F$ is a mapping $f : U \rightarrow X(f)$). A member f in F is called a factor, and $\mathbf{0}$ is called empty factor. The operations \vee and \wedge are called the synthesis and decomposition of factors respectively.

We often limit the definition that $F_o = \{f_1, \dots, f_n\} = f_1 \vee \dots \vee f_n$, and $F = P(F_o)$, F is the power of F_o . In this case, the operations \vee and \wedge become to \cup and \cap respectively.

This definition is a relaxed vision given here. Originally, the index set was defined as a Boolean algebra $F = (F; \vee, \wedge, \neg, \mathbf{0}, \mathbf{1})$ [26], which is too strong for our study. Of course, we often treat F as a Boolean algebra if it is needed.

Remark 3. About the name "Factors Space", we need to emphasize here: our definition is different from the factor space $X_{/\sim}$ representing the set of classes classified by the equivalent relation \sim . From now on, we use the word Factors, not Factor, in writings; But, we keep using the word in factor analysis initiated by L. L. Thurston [15]. As the great psychologist in 1931, even though he did not take high mathematical description into cognitive psychological measurement, he raised the banner of factors more than half a century earlier as our pioneer. Today, factors space inherits his target forward.

Factors space is a generalization of Cartesian coordinate space. Any quantitative Cartesian space can be viewed as a factors space since every variable is a factor; factors space extends those coordinate spaces from fixed dimension to variable dimension. The main breakthrough of factors space is that the states of factor can be changed from quantitative to qualitative. There are two types of states of a factor; one is quantitative; for example, the state of height can be measured as a real number in the interval $[10, 250]$ (cm); the other type is qualitative; the state is taken as a word in natural language, tall, middle or short. In most cases, the two types of states can be employed simultaneously, and form a pair of state space $X(f)$ and $\underline{X}(f)$, the former one stands for qualitative and the latter one the quantitative State space. In the AI applications, we employ qualitative on semantic description and quantitative on formulae calculation. The most important point is how to transfer each other. Fuzzy sets theory provides technique to do such a kind of transformations, which enables FS to build its coordinate system by available axes.

4 Background set and correlation distribution

In the Section, we need to introduce the core concept of factors space.

Definition 4. Given a factors space $\psi = (U, X_F)$ with $F = P(F_o)$ and $F_o = f_1 \vee \dots \vee f_n$, denote

$$R = F_o(U) = \{(x_1, \dots, x_n) | \exists u \in U; x_1 = f_1(u), \dots, x_n = f_n(u)\}. \quad (5)$$

Which is called the background set of ψ .

The background set R is the image of universe U in $X(F_o)$, which can be understood the real Cartesian product space of f_1, \dots, f_n . Outside of R , the state-configuration of the n factors is not possible to be realized, so that R can be seen as the state space of the combined factor F_o . When f_1, \dots, f_n are independent, we have that $R = X(f_1) \times \dots \times X(f_n)$; otherwise, there are correlations existing between those factors. The background relation R is also called the correlation of factors f_1, \dots, f_n .

Correlation of factors is the source of factors space theory, which generates causality, logic, concept generation and decision making. Correlation of factors is the foundation of cognition math. A basic theorem was a proposal in paper [23].

We give an intuitive proof as follows:

Theorem 5. *Let R be the correlation of f and g , set $X = X(f)$ and $Y = X(g)$. For any $E \subseteq X$ and $E' \subseteq Y$, the inference "If x is in E then y is in E' " is identically true if and only*

$$(E \times Y) \cap R \subseteq (X \times E') \cap R. \quad (6)$$

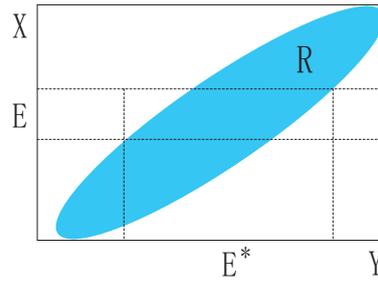


Figure 1: Background relation determines inference

Proof: The proposition " x is in E " is equivalent to " x is in E and y is in Y " and " (x, y) is in $E \times Y$ ". Since (x, y) is in R always, The proposition " x is in E " is equivalent to " (x, y) is in $(E \times Y) \cap R$ ". Similarly, the proposition " y is in E' " is equivalent to " (x, y) is in $(X \times E') \cap R$ ". The inference "If x is in E then y is in E' " is true if and only $(E \times Y) \cap R \subseteq (X \times E') \cap R$ \square

Since the factors f and g can be complex factors in high dimension space, theorem 5 can be used in general cases, which tells us that the correlation of factors determines all inferences between those factors.

It is obvious that the background set is indeed the formal background in FCA. FS theory further underscoring the significance of the formal background emphasized by Wille.

So far, we get background set as a theoretical base with determinacy. But, it has to be extended to handle uncertainty from the following two reasons:

1) An object u in U is not a point but a granule, which may be a complex system. According to the viewpoint of granular computing [33], the size of granule can be varying in hierarchical model. For example, u was a man, the factor $f = height$ maps u to a determinate state $f(u) = Tall$ in $X(f)$, but, when the size of granule u is changed from a man to a group of men, what is the state $f(u)$? Here comes the uncertainty.

2) U is a theoretical term, which is not available in practice. For example, consider that U consists of all people. Where to identify a person? In the East or the West? When to identify a person? Include those who have been died or will be born? In practice, the universe is a sampling of U with randomness.

From the two reasons, we need to treat u as random variable and extend background set to background distribution relatedly.

Definition 6. Given a factors space $\psi = (U, X_F)$ with $F = P(F_o)$ and $F_o = f_1 \vee \dots \vee f_n$, and suppose that $X(f_i) = \{a_{i1}, \dots, a_{ik(i)}\}$ ($i = 1, \dots, n$) are qualitative state spaces both. Let p be the probability carried by factors from random variable u . Denote

$$r(x_1, \dots, x_n) = p(f_1(u) = x_1, \dots, f_n(u) = x_n) \quad (\sum \{b(x_1, \dots, x_n) | x_i = a_{i1}, \dots, a_{ik(i)} \ (i = 1, \dots, n)\} = 1) \quad (7)$$

where r is called the background distribution of ψ , or the correlation distribution between factors f_1, \dots, f_n .

The definition can be extended to quantitative state spaces.

5 Factors space in concept generation

Concept generation is a key point and a basic topic in cognition math. We only introduce the topic on concept generation in the paper.

5.1 General target of AI

Suppose that there are some intelligent problems about bodies of water treated by hydrologists. There are 21 objects as follows: 1. Tarn, 2. Trickle, 3. Rill, 4. Beck, 5. Rivulet, 6. Runnel, 7. Brook, 8. Bum, 9. Stream, 10. Torrent, 11. River, 12. Canal, 13. Lagoon, 14. Lake, 15. Mere, 16. Splash, 17. Pond, 18. Pool, 19. Puddle, 20. Reservoir, 21. Sea. We take those terms as objects to form a set U . The tasks of AI are treating objects according to a general target: Picking up information and transferring to the knowledge. Even though there are only 21 objects, each of them is not a piece of body in the land, but has been classified as a word in hydrology, and they are rambling too. The first task of AI is organizing objects into a conceptual system: From touse to intellect.

5.2 Using factor to distinguish and classify objects

Concepts come from comparing and distinguish objects; it could not be done without factors.

Definition 7. We call object u and v in U are different if there is a factor f defined on U such that

$$f(u) \neq f(v). \quad (8)$$

Otherwise, they are called the same with respect to factor f .

Any two things in the world have the same points and different points simultaneously, without factors, we couldn't judge if two things are same or different.

A factor f is a mapping defined on the universe U , which determines an equivalent relation \sim on U :

$$u \sim v \text{ if and only if } f(u) = f(v).$$

The equivalent relation determines a classification on U ; if we consider only qualitative state space here, then the classified factorial space can be denoted as

$$U_{/f} = \{C_k = (u_{k1}, \dots, u_{kn(k)})\}_{k=1, \dots, K} \quad (9)$$

where the number of states of $X(f)$ is $m = n(1) + \dots + n(K)$, $n(j)$ is the number of objects in k^{th} subclass.

Definition 8. [38] Denote

$$c_f = n(1)(n(1) - 1) + \dots + n(K)(n(K) - 1) / m(m - 1) \quad (10)$$

Which is called the distinguish degree of factor f with respect to U .

The bigger the distinguish degree, the finer the division taken by f on U .

Table 1: Factorial table for bodies of water

U	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
f_1	H	N	N	N	N	N	N	N	N	N	N	A	N	N	A	N	A	N	N	A	N
f_2	M	R	R	R	R	R	R	R	R	R	R	R	M	M	M	M	M	M	M	M	M
f_3	S	S	S	S	S	S	S	S	S	S	S	S	I	S	S	S	S	S	S	S	I
f_4	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	T	C	C

5.3 Atomic concepts

As for the bodies of water, to pick up information and organize objects into a conceptual system, we need to consider some factors: Is the body natural or artificial? How is its manner, running or muddy? Is the water body in sea or in land? Is it constant or temporary? And so on. Those factors are often the names of attributes, but factors do not play the role of passive descriptors but initiative leaders. In the bodies of water, we can define that $f_1 = \text{Natural?}$, $X(f_1) = \{\text{Natural}, \text{Artificial}\}$; $f_2 = \text{Running?}$, $X(f_2) = \{\text{Running}, \text{Muddy}\}$; $f_3 = \text{Sea – inland}$, $X(f_3) = \{\text{Sea}, \text{Inland}\}$; $f_4 = \text{Timeliness}$, $X(f_4) = \{\text{Constant}, \text{Temporary}\}$. Then we can get a factor space $\psi = (U, X_F)$ with $F_o = f_1 \vee f_2 \vee f_3 \vee f_4$, The factor tableau can be taken as the sampling in $X = X(f_1) \times X(f_2) \times X(f_3) \times X(f_4)$. Citing the data from paper [5], we get Table 1.

How to analyze on the tableau? Given a factors space $\psi = (U, X_F)$ with $F_o = f_1 \vee \dots \vee f_n$, for any $[\mathbf{a}] \in X(F_o)$, denote

$$[\mathbf{a}] = F_o^{-1}(\mathbf{a}) = \{u | F_o(u) = (f_1(u), \dots, f_n(u)) \in \mathbf{a}\} \tag{11}$$

If $[\mathbf{a}] \in R$, i.e., $[\mathbf{a}] \neq \emptyset$, we say that \mathbf{a} is an atomic intension, or say a is a real configuration. The collection of all the atom intensions form the division $U_{/F_o} = \{[\mathbf{a}]\}$, which includes all subclasses divided by factors f_1, \dots, f_n . It is clear that F_o is an isomorphic mapping from $U_{/F_o}$ to R .

Definition 9. Given factors space $\psi = (U, X_F)$ with $F = P(F_o)$ and $F_o = f_1 \vee \dots \vee f_n$, for any $\mathbf{a} \in R$, the pair $\alpha = (\mathbf{a}, [\mathbf{a}])$ is called an atomic concept and $[\mathbf{a}]$, \mathbf{a} are called the extension and intension of α respectively.

Return to the bodies of water, the Cartesian product space $X(f_1) \times X(f_2) \times X(f_3) \times X(f_4)$ concludes $2 \times 2 \times 2 \times 2 = 16$ elements, but, the real product X_F , i.e., the background set R includes 6 elements only

$$R = \{\text{NMSC}, \text{NRSC}, \text{ARSC}, \text{NMIC}, \text{AMSC}, \text{NMST}\}$$

According to our definition, there are six atomic concepts:

$$\begin{aligned} \alpha_1 &= (\text{NMSC}, \{1,14,18\}), \alpha_2 = (\text{NRSC}, \{2,3,4,5,6,7,8,9,10,11\}), \alpha_3 = (\text{ARSC}, \{12\}), \\ \alpha_4 &= (\text{NMIC}, \{13,21\}), \alpha_5 = (\text{AMSC}, \{15,17,20\}), \alpha_6 = (\text{NMST}, \{16,19\}). \end{aligned}$$

5.4 How important are the atomic concepts?

Atomic concepts have 4 basic functions in AI:

1. Atomic concepts provides the finest classification of U under factors f_1, \dots, f_n .

Theorem 10. Given factors space $\psi = (U, X_F)$ with $F = P(F_o)$ and $F_o = f_1 \vee \dots \vee f_n$, each atomic concept $\alpha = (\mathbf{a}, [\mathbf{a}])$ is not able to be refined (i.e., $[\mathbf{a}]$ is not able to be divided) by the factors $f_1 \vee \dots \vee f_n$.

Proof: For any atomic extension $[\mathbf{a}]$ and any $u, v \in [\mathbf{a}]$, we have that $f_i(u) = [\mathbf{a}] = f_i(v)$ for $i = 1, \dots, n$, so that $u \sim v$ for $i = 1, \dots, n$. \square

The smaller the number of atomic concepts with respect to the number of objects in U , the simpler and clearer the structure of knowledge is but the rougher the classification is. To break the limit of finest division, we need to find out more factors, and this is a deepening process and will be discussed in the future.

2. The finest classification provides the universal code for knowledge organization, retrieval and inquiring. Of course, the code is edited under the guidance of factors.

3. An answer system can answer following questions based on atomic concepts:

Is Rill α_1 ? No, since $\alpha_1 = (\text{NMSC}, \{1,14,18\})$ does not includes Rill (3).

Does α_2 include Rill? Yes, since $\alpha_2 = (\text{NRSC}, \{2,3,4,5,6,7,8,9,10,11\})$ includes Rill.

Does Rill have the attribute A ? No, since $fi(\text{Rill}) = A \neq N$.

What is the concept an object with attributes N and S belonging to? It may α_1, α_2 or α_4 be.

4. All atomic concepts form the background relation or the correlation R . What is the shape of R ? What is the causality generated from R ? There comes a lot of functions applied in AI, we will discuss in the future.

5.5 Basic concepts semi-lattice

Definition 11. Given a factors space $\psi = (U, X_F)$ with $F = P(F_o)$ and $F_o = f_1 \vee \dots \vee f_n$, for any $A \subseteq R$, denote $[A] = \cup\{[\mathbf{a}] | \mathbf{a} \in A\}$, For any $A \subseteq R$, $\gamma = (A, [A])$ is called a concept with the intension A and extension $[A]$. Set $\Gamma = \{\gamma = (A, [A]) | A \subseteq R\}$, $\Gamma = (\Gamma, \text{OR}, \text{AND})$ is called the conceptual lattice on $\psi = (U, X_F)$, Where $\gamma \text{OR} \gamma' = (A \text{OR} A', [A] \cup [A'])$ and $\gamma \text{AND} \gamma' = (A \text{AND} A', [A] \cap [A'])$; even $\Gamma = (\Gamma, \text{OR}, \text{AND}, \text{NOT})$ is called the conceptual Boolean algebra on $\psi = (U, X_F)$, Where $\text{NOT} \gamma = (\text{NOT} A, [A]^c)$.

Do not be afraid that the generated concepts are too little; we are afraid that too much. If there are k atomic concepts, then 2^k new concepts will be generated. To avoid the information overflow, we generate basic concepts only.

Definition 12. A concept is called basic if its intension can be stated in a conjunctive normal form.

Intuitively speaking, the extension of a basic concept can be mapped to a hyper-rectangle within R in $X(F_o)$.

In bodies of water, each atomic concept is a basic concept. The concept $\gamma = (\text{NMC}, \{1,14,18,13,21\})$ is a basic concept since its intension can be written to the conjunctive normal form $N \wedge M \wedge (S \vee I) \wedge C = N \wedge M \wedge X(f_3) \wedge C = N \wedge M \wedge C$.

How to get the basic concept semi-lattice? We introduce Wang's algorithm [21] by the example of bodies of water.

1) Calculating distinguish degree c_f of each factor and then changing the order of columns to do f -classification on U , where f holds the maximal distinguish degree c_f .

$$m = 22$$

$$f_1: n(\text{N}) = 16, n(\text{A}) = 5, c_{f_1} = 1 - (16 \times 15 + 5 \times 4) / 21 \times 20 = 0.4;$$

$$f_2: n(\text{R}) = 11, n(\text{M}) = 10, c_{f_2} = 1 - (11 \times 10 + 10 \times 9) / 21 \times 20 = 0.5;$$

$$f_3: n(\text{S}) = 19, n(\text{I}) = 2, c_{f_3} = 1 - (19 \times 18 + 2 \times 1) / 21 \times 20 = 0.2;$$

$$f_4: n(\text{C}) = 19, n(\text{T}) = 2, c_{f_4} = 0.2;$$

$$c_{f_2} > c_{f_1} > c_{f_3} = c_{f_4}.$$

U	12	2	3	4	5	6	7	8	9	10	11	1	13	14	15	16	17	18	19	20	21
f_1	A	N	N	N	N	N	N	N	N	N	N	H	N	N	A	N	A	N	N	A	N
f_2	R	R	R	R	R	R	R	R	R	R	R	M	M	M	M	M	M	M	M	M	M
f_3	S	S	S	S	S	S	S	S	S	S	S	S	I	S	S	S	S	S	S	S	I
f_4	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	T	C	C

We get the f_2 -classification $U = U_1\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} + U_2\{1, 13, 14, 15, 16, 17, 18, 19, 20, 21\}$.

Repeatedly, do classification on U_1 and return to step 1):

U	12	2	3	4	5	6	7	8	9	10	11
f_1	A	N	N	N	N	N	N	N	N	N	N
f_2	R	R	R	R	R	R	R	R	R	R	R
f_3	S	S	S	S	S	S	S	S	S	S	S
f_4	C	C	C	C	C	C	C	C	C	C	C

$m=11$

$$f_1 : n(N) = 10, n(A) = 1, c_{f_1} = 1 - (10 \times 9 + 1 \times 0) / 11 \times 10 = 9/11;$$

$$f_2 : n(R) = 11, n(M) = 0, c_{f_2} = 1 - (11 \times 10) / 11 \times 10 = 0;$$

$$f_3 : n(S) = 11, n(I) = 0, c_{f_3} = 1 - (11 \times 10) / 11 \times 10 = 0;$$

$$f_4 : n(C) = 11, n(T) = 0, c_{f_4} = 1 - (11 \times 10) / 11 \times 10 = 0.$$

$$c_{f_1} > c_{f_2} = c_{f_3} = c_{f_4}$$

We get the f_1 -classification $U_1 = U_{11}\{12\} + U_{12}\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$.

Similarly, we can repeat the step 1) on other sub-classes. Finally, we get a tree with following branches

$$U_2 = U_{21}\{1, 13, 14, 16, 18, 19, 21\} + U_{22}\{15, 17, 20\};$$

$$U_{21} = U_{211}\{13, 21\} + U_{212}\{1, 14, 16, 18, 19\};$$

$$U_{212} = U_{2121}\{1, 14, 18\} + U_{2122}\{16, 19\}.$$

The figure of the tree is drawn in Fig.2.

The complexity of basic concept semi-lattice is $O(m^2n)$, where m and n are the number of rows and columns of factorial tableau respectively.

In Fig.2, each node is a basic concept, and the node at the top is the original concept with extension U . When a baby born, the origin is zero concept (Empty, U), No intension statement, the extension is the chaotic universe. From top to bottom, the intension increases and the extension shrinks. The undermost node is not a concept but an ideal extreme with empty extension. We add it in the Fig.2 to match the original figure drawn by Ganter and Wille.

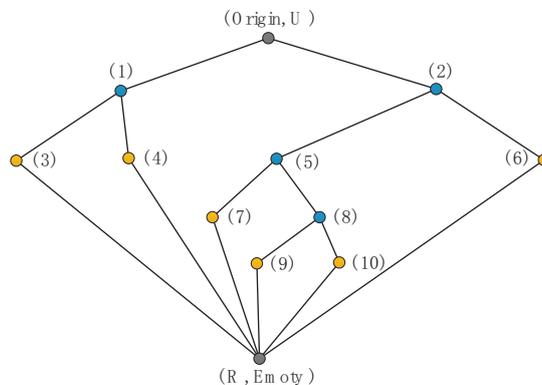


Figure 2: Semi-lattice of basic concepts

Theorem 13. *The operation OR may not be closed for two basic concepts except that the intension of two concepts are different at only one factor.*

Proof is obvious, we only hint that $\gamma_1 \text{ OR } \gamma_4 = (\text{NMSCORNMIC}, \{1, 14, 18\} \cup \{13, 21\}) = (\text{NMC}, \{1, 14, 18, 13, 21\})$, which is the basic concept γ mentioned above. Performing the operation OR on two basic concepts, we get a basic concept. What is the condition of the result? The intensions of the two basic concepts are NMSC and NMIC, and their difference is on the factor f_3 only. Without the condition, the operation of OR is not closed in basic concepts. Be careful, all basic concepts form a semi-lattice, but may not be a lattice.

5.6 The consistence with formal concept analysis

Concept generation was established by Wille, who defined the formal background as $K = (G, M, I)$, where G is a group of objects, M a group of attribute values, I a relation from G to M : $(g, m) \in I$ implies that object g holds attribute value m . For any $A \subseteq G$, define $s(A) = \{m \in M | \forall g \in A; (g, m) \in I\}$, and for any $B \subseteq M$, define $t(B) = \{g \in G | \forall m \in B; (g, m) \in I\}$. $\alpha = (B, A)$ is called a concept if the following convolution principle is holds:

$$s(A) = B, \quad t(B) = A \quad (12)$$

Theorem 14. *Given a factors space $\psi = (U, X_F)$ with $F = P(F_o)$ and $F_o = f_1 \vee \dots \vee f_n$, for any basic concept $\alpha = (A, [A])$, $[A]$ and A satisfy the convolution principle.*

Proof: Let $\alpha = (\mathbf{a}, [\mathbf{a}])$ be an atomic concept. Taking $G = U$, $M = X(F_o)$ and $I = \{(u, \mathbf{a}) | \mathbf{a} = F_o(u)\}$, we have

$$\begin{aligned} s([\mathbf{a}]) &= \{m \in M | \forall g \in [\mathbf{a}]; (g, m) \in I\} = \{F_o(u) | u \in [\mathbf{a}]\} = \mathbf{a}; \\ t([\mathbf{a}]) &= \{u \in U | (u, \mathbf{a}) \in I\} = \{u \in U | F_o(u) = \mathbf{a}\} = [\mathbf{a}]. \end{aligned}$$

Hence $[\mathbf{a}]$ and \mathbf{a} satisfy convolution Eq.(12).

Let $\alpha = (A, [A])$ be a basic concept. We have

$$\begin{aligned} s([A]) &= \{A \subseteq X(F_o) | \forall u \in [A]; F_o(u) \in A = \mathbf{a}\}; \\ t(A) &= \{u \in U | F_o(u) \in A\} = [A]. \end{aligned}$$

Hence $[A]$ and A satisfy convolution Eq.(12). □

The proposition proves that the definition of basic concepts in FS is consistent with the original definition in FCA. It does also hint that general concepts may not satisfy the convolution principle.

What is the relationship between Table 1 and the original Table of Wille? The original table shows a formal background with 22 objects and 8 attribute values. Table 1 replaces the 8 attribute values by 4 factors, and since the object *channel* has no attribute value in the original table, it has been taken out of the table.

As shown in Fig.2, apart from the top and bottom nodes, there are 8 red atomic and 4 blue non-atomic basic concepts. To be easy to match to Fig.3, we number the nodes as follows:

Node	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$[A]$	U_1	U_2	U_{11}	U_{12}	U_{21}	U_{22}	U_{211}	U_{212}	U_{2121}	U_{2122}

It is obvious that nodes (3),(4),(6),(7),(9),(10) are atomic, color red, nodes (1),(2),(5),(8) are non-atomic basic concepts, color blue. We can see that Fig.2 is consistent with Fig.3. Indeed, the semi-lattice drawn in Fig.2 is a sub-lattice drawn in Fig.3.

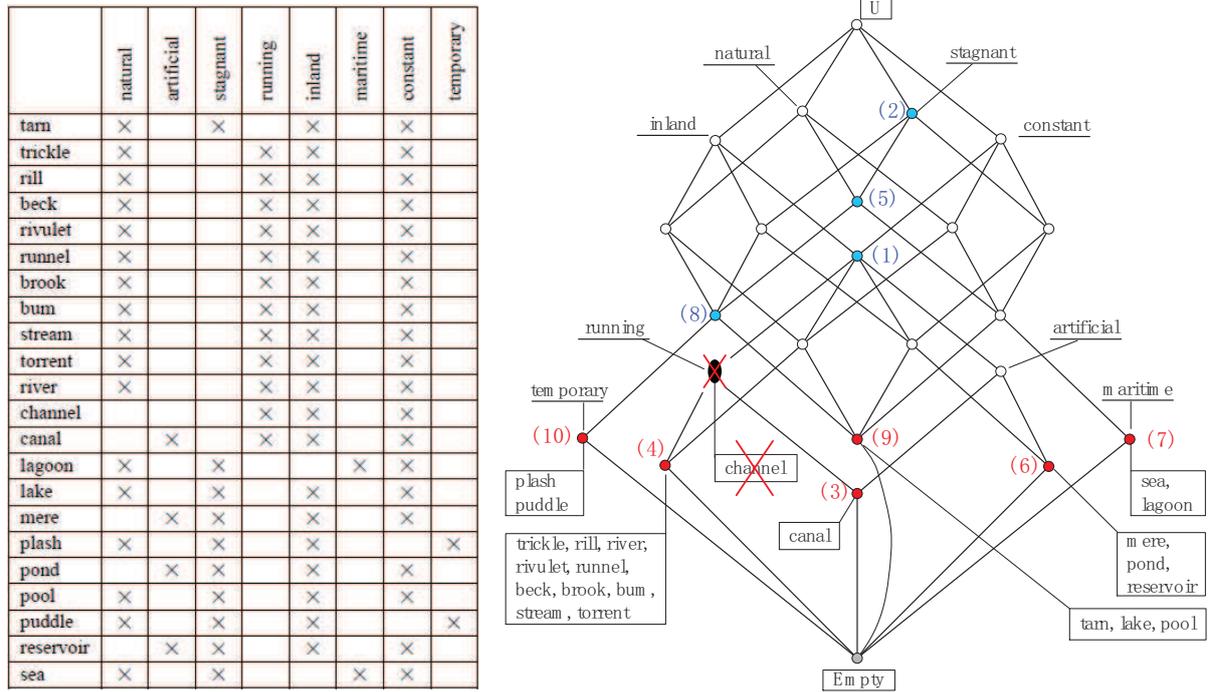


Figure 3: Conceptual lattice(Ganter & Wille [5]).

In practical applications, the generated basic concepts are too much, we need to review those concepts by experts, rarely select those concepts who can be understood by human being, and then rename them into a special database. Fig.2 decreases the number of non-atomic basic concepts from 15 in origin to 4 in Fig.2.

However, we can change the semi-lattice by changing the order of factor-divisions, and the union of that semi lattice can completely full the original 'concept lattice':

Theorem 15. *Given a factors space $\psi = (U, X_F)$ with $F = P(F_o)$ and $F_o = f_1 \vee \dots \vee f_n$, for any basic concept $\gamma = (A, [A])$, there is a semi-lattice S generated by Wang's algorithm taking γ as a member in it.*

Proof: Let the intension A is written in the conjunction normal form: $A = A_{i(1)} \wedge \dots \wedge A_{i(k)}$ ($k \leq n$), where $A_{i(j)} = a_{i(j),j(1)} \vee \dots \vee a_{i(j),j(r)}$ is a disjunction of atomic intensions with respect to the factor $f_{i(j)}(j = 1, \dots, k)$. According to Wang's algorithm, doing classifications by factors $f_i(1), \dots, f_i(k)$ successively, the basic concept must be occurred in the semi-lattice of basic concept. □

Basic concepts have 2 basic functions in AI:

1. A set of inference rules can be shown in the graph of semi-lattice.

Let S be the graph of a semi-lattice. From bottom to up, any node $\gamma = (A, [A])$ links at least one node $\gamma' = (A', [A'])$. It is obvious that the inference arrow $\gamma \Rightarrow \gamma'$ is constant true.

2. An answer system can answer the following questions based on basic concepts:

Is Pool in U_1 ? No, since Pool is α_1 numbered red 9, which links to blue 8, then links to blue 5, then links to blue 2, no way to link to blue 1. It does not belong to U_1 .

Is Drill in U_1 ? Yes, since Drill is α_2 numbered red 4, which links to the node numbered blue 1, having extension U_1 .

6 Conclusion

Factor space provides a new view point toward formal concept analysis, which promotes the signification of formal background. By means of the support of FS, all problems in FCA or RS can get more clear and simple statement, and the N-hard trap FCA and RS encountered will be overcome by faster algorithms. Concept generation is a key point and basic problem in artificial intelligence and factor space has got good progress on concept generation, which will be an important tool for cognition, information and data science.

Acknowledgement

The authors specially thank Professor P.Z. Wang, was a good friend and collaborator with the father of fuzzy sets: Lotfi A. Zadeh (1921-2017), for his guidance and valuable suggestions. This study was partially supported by the grants (Grant Nos. 61350003, 11401284, 70621001, 70531040) from the Natural Science Foundation of China, and the grant (Grant Nos. L2014133) from the department of education of Liaoning Province.

Bibliography

- [1] Cheng, Q.F.; Wang, T.T.; Guo, S.C.; Zhang, D.Y.; Jing, K.; Feng, L.; Wang, P.Z. (2017); The Logistic Regression from the Viewpoint of the Factor Space Theory, *International Journal of Computers Communications & Control*, 12(4), 492–502, 2017.
- [2] Cui, T.J.; Wang, P.Z.; Ma, Y.D. (2016); Structured representation methods for 01 space fault tree, *J Dalian Jiaotong Univ*, 37(1), 82–87, 2016.
- [3] Dzitac, I. (2015), The Fuzzification of Classical Structures: A General View, *International Journal of Computers Communications & Control*, 10(6), 772-788, 2015.
- [4] Dzitac, I.; Filip, F.G. ; Manolescu, M.J. (2017); Fuzzy Logic Is Not Fuzzy: World-renowned Computer Scientist Lotfi A. Zadeh, *International Journal of Computers Communications & Control*, 12(6), 748-789, 2017.
- [5] Ganter, B.; Wille, R. (1996); Formal concept analysis, *Wissenschaftliche Zeitschrift-Technischen Universitat Dresden*, 45, 8–13, 1999.
- [6] Kandel, A.; Peng, X.T.; Cao, Z.Q.; Wang, P.Z. (1990); Representation of concepts by factor spaces, *Cybernetics and Systems: An International Journal*, 21(1), 43–57, 1990.
- [7] Li, H.X.; Wang P.Z.; Yen, V.C. (1998); Factor spaces theory and its applications to fuzzy information processing.(I). The basics of factor spaces, *Fuzzy Sets and Systems*, 95(2), 147–160, 1998.
- [8] Li, H.X.; Yen, V.C.; Lee, E.S. (2000); Factor space theory in fuzzy information processing-Composition of states of factors and multifactorial decision making, *Computers & Mathematics with Applications*, 39(1), 245–265, 2000.
- [9] Li, H.X.; Yen, V.C.; Lee, E.S. (2000); Models of neurons based on factor space, *Computers & Mathematics with Applications*, 39(12), 91–100, 2000.

- [10] Li, H.X.; Chen, C.P.; Yen, V.C.; Lee, E.S. (2000); Factor spaces theory and its applications to fuzzy information processing: Two kinds of factor space canes, *Computers & Mathematics with Applications*, 40(6-7), 835–843, 2000.
- [11] Li, H.X.; Chen, C.P.; Lee, E.S. (2000); Factor space theory and fuzzy information processing-Fuzzy decision making based on the concepts of feedback extension, *Computers & Mathematics with Applications*, 40(6-7), 845–864, 2000.
- [12] Liu, H.T.; Guo, S.C. (2015); Inference model of causality analysis, *Journal of Liaoning Technical University(Natural Science)*, 2015, 34(1), 124–128.
- [13] Liu, Z.L. (1990); *Factorial Neural Networks*, Beijing Normal University Press, 1990.
- [14] Pawlak, Z. (1982); Rough sets, *International Journal of Parallel Programming*, 11(5), 341–356, 1982.
- [15] Thurstone L. L. (1931); Multiple factor analysis, *Psychological Review*, 38(5), 406–427, 1931.
- [16] Vesselenyi, T.; Dzitac, I.; Dzitac, S.; Vaida, V. (2008); Surface roughness image analysis using quasi-fractal characteristics and fuzzy clustering methods, *International Journal of Computers Communications & Control*, 3(3), 304–316, 2008.
- [17] Wang, P.Z. (1981); Randomness, *Advance of Statistical Physics*, Science and Technology Press, 1981.
- [18] Wang, P.Z. (1985); *Fuzzy sets and falling shadows of random set*, Beijing Normal University Press, 1985.
- [19] Wang, P.Z. (1990); A factor spaces approach to knowledge representation, *Fuzzy Sets and Systems*, 36(1), 113–124, 1990.
- [20] Wang, P.Z. (1992); Factor space and concept description, *Journal of Software*, 1, 30–40, 1992.
- [21] Wang, P.Z. (2013); Factor spaces and factor data-bases, *Journal of Liaoning Technical University (Natural Science)*, 32(10), 1–8, 2013.
- [22] Wang, P.Z. (2015); Factors space and data science, *Journal of Liaoning Technical University (Natural Science)*, 34(2), 273–280, 2015.
- [23] Wang, P.Z.; Guo, S.C.; Bao, Y.K.; Liu, H.T. (2014); Causality analysis in factor spaces, *Journal of Liaoning Technical University (Natural Science)*, 33(7), 1–6, 2015.
- [24] Wang, P.Z.; Jiang, A. (2002); Rules detecting and rules-data mutual enhancement based on factors space theory, *International Journal of Information Technology & Decision Making*, 1(01), 73–90, 2002.
- [25] Wang, P.Z.; Li, H.X. (1995); *Fuzzy system theory and fuzzy computer*, Publishing Company of Science, 1995.
- [26] Wang, P.Z.; Li, H.X. (1994); *A mathematical theory on knowledge representation*, Tianjin Scientific and Technical Press, 1994.
- [27] Wang, P.Z.; Liu, Z.L.; Shi, Y.; Guo, S.C. (2014); Factor space, the theoretical base of data science, *Annals of Data Science*, 1(2), 233–251, 2014.

- [28] Wang, P.Z.; Ouyang, H.; Zhong, Y.X.; He, H.C. (2016); Cognition math based on factor space, *Annals of Data Science*, 3(3), 281–303, 2016.
- [29] Wang, P.Z., Sugeno, M. (1982); The factor fields and background structure for fuzzy subsets, *Fuzzy Mathematics*, 2(2), 45–54, 1982.
- [30] Wang, H.D.; Wang, P.Z.; Shi, Y.; Liu, H.T. (2014); Improved factorial analysis algorithm in factor spaces, *International Conference on Informatics*, 201–204, 2014.
- [31] Wang, P.Z.; Zhang, X.H.; Lui, H.C.; Zhang, H.M., Xu, W. (1995); Mathematical theory of truth-valued flow inference, *Fuzzy Sets and Systems*, 72(2), 221–238, 1995.
- [32] Wille, R. (1982); Restructuring lattice theory: an approach based on hierarchies of concepts, *Ordered sets*, Springer Netherlands, 445–470, 1982.
- [33] Yao, Y. (2009); Three-Way Decision: An Interpretation of Rules in Rough Set Theory, *RSKT*, 9, 642–649, 2009.
- [34] Yuan, X.H.; Wang, P.Z.; Lee, E.S. (1992); Factor space and its algebraic representation theory, *J Math Anal Appl.*, 171(1), 256–276, 1992.
- [35] Yuan, X.H.; Wang, P.Z.; Lee, E.S. (1994); Factor Rattans, Category FR (Y), and Factor Space, *Journal of Mathematical Analysis and Applications*, 186(1), 254–264, 1994.
- [36] Zadeh, L.A. (1965); Fuzzy sets, *Information and control*, 8(3), 338–353, 1965.
- [37] Zeng, F.H.; Zheng, L. (2017); Sample cultivation in Factorial analysis, *Journal of Liaoning Technical University (Natural Science)*, 36(3), 320–323, 2017.
- [38] Zeng, F.H.; Li, Y. (2017); An improved decision tree algorithm based on factor space theory, *Journal of Liaoning Technical University (Natural Science)*, 36(3), 109–112, 2017.