Semilattices Category and Data Representations for Algebraic Machine Learning

Dmitry V. Vinogradov[1]

[1]*Dorodnicyn Computing Center FRC Computer Science and Control, 40 Vavilova Str., Moscow, 119333, Russia*

Abstract

Category of semilattices coincides with algebras on a corresponding monad. This statement implies Cartesian-closeness of this category. It means applicability of well-known constructions of (direct) limits of finite functors (such as Cartesian products, equalizers, amalgams, etc.). These structures are useful for representation complex attributes to describe training and test examples for algebraic (lattice-theoretic) machine learning. The paper contains an exposition of these ideas to ‘old-fashioned’ machine learning specialists.

Keywords

semilattice, monad, category of algebras, limit of functor, FCA, entropy

1. Introduction

In Russia there exists a research group leading by Prof. V.K. Finn aimed to investigation of Knowledge Discovery schemes. Originally, these researchers used Boolean algebra and many-valued logic's means. Result of the research was the theory called ‘JSM method’ of automatic hypotheses generation [1]. Modern approach applies lattice-theoretic means from Formal Concept Analysis (FCA) [2,3] to extend the theory on more general situation of binary similarity operation defined between arbitrary objects. This extension also led to more efficient algorithms for Inductive Generalization procedure (Induction).

However JSM method approach has several drawbacks: high computational (exponential memory) complexity in worst case (of Boolean algebra), and over-training phenomena [4]. The author develops probabilistic approach to avoid these obstacles [5]. This theory named ‘VKF method’ in honor of V.K. Finn.

Fundamental Theorem of FCA implies that every training sample can be represented as list of ‘bitsets’ – strings of bits of fixed length – together with bit-wise multiplication as the similarity operation. Hence the prominent problem of VKF method is to encode training and test examples by bitsets. The author applied modern proof of Fundamental Theorem of FCA to encode objects described by attributes with discrete values that form semilattice [6]. Then the author [7] used an analogue of J.R. Quinlan's approach in C4.5 decision tree algorithm [8] to objects with continuous features. Initially, the procedure splits the entire domain of a continuous attribute into several intervals in order to achieve a minimum mean entropy. Then it uses the generated thresholds to encode the attribute value that falls within one of these intervals. Currently, the author and his PhD student L.A. Yakimova are investigating the possibility of using of Sparse Autoencoder (a special kind of neural network) to encode images by sparse bitsets.

Usage of several representations of attributes of objects generates a problem of comparing results of experiments with differently represented training objects with fixed number of features. Same problem occurs when descriptions of training and test examples expand by additional attributes. Finally, there exist a sequential VKF method, where partially (with respect to some subset of attributes) defined hypotheses can be extended by taking into account additional features.

[1]Russian Conference on Artificial Intelligence (RCAI-2021), October 11–16, 2021, Taganrog, Russia
EMAIL: vinogradov.d.w@gmail.com
ORCID: 0000-0001-5761-4706

Initially, idea of extending of training sample went from JSM method. However, it adds new training examples and assumes that all attributes that describe objects are observed simultaneously. This assumption is hardly adequate in practice, when a decision maker receives information sequentially, in blocks. Many alternative paradigms, that is, decision trees, recurrent neural networks, and probabilistic automata, take this sequential process in account. Sequential VKF method tries to remove this restriction of JSM method too.

This paper presents a category theoretic representation for a sequential version of the VKF method, in which a semilattice of object descriptions is mapped to a Cartesian product of semilattices corresponding to the values of attributes that describe training examples and the similarities generated from them.

A comparison of hypotheses with different features representations that coincide on some common set of features corresponds to a well-known construction of amalgam of two algebras with respect to homomorphisms into third one.

Both above constructions are partial cases of direct limits in category of semilattices. Hence, it is clear that category-theoretic language is best choice for description of such constructions. Moreover, it turned out that the category of algebras over the corresponding monad admits a pair of conjugate functors, one of which generates free objects in this category. And most importantly, the conjugation isomorphism defines the polar operation (or ‘the global similarity’ in the terminology of the JSM method), which is the main tool of both VKF method and JSM method.

To understand the paper readers need to know basic notions of Formal Concept Analysis [2]. No knowledge of category theory is assumed. The author recall the main ideas of this theory in accordance with the fundamental book [9] below. The paper contains only proofs of 2 fundamental results: coincidence of the category of semilattices with the category of algebras over monad with power-set functor and singleton natural transformation and explicit construction of polar (‘global similarity’ in JSM language) through conjugation of forgetting and (free-algebra) generating functors.

2. Bitset Representations

Recall, that ‘bitset’ is a string of bits of fixed length. We consider bit-wise multiplication of two bitsets as ‘local similarity’ between them. This representation is computer-oriented, because

1. there exist data types in modern programming languages (for instance, boost::dynamic_bitset for C++);
2. bit-wise multiplication is computational effective (on GPGPU it consumes 4 ticks, on CPU – only 1 tick);
3. modern compilers admit various optimizations on bitsets (for example, vector-parallelism).

VKF method admits more general situation: it needs semilattice – a set X (of ‘fragments’) together with binary operation (called ‘local similarity’ or ‘intersection’). This operation must be idempotent, commutative and associative

However, Fundamental Theorem of FCA see [2] states that every finite semilattice can be reconstructed (up to isomorphism) as a lattice of ‘concepts’ of training sample, where sample is a list of bitsets, with help of bit-wise multiplication. VKF method [5] renames a concept into ‘candidate into hypothesis’ (or simply ‘candidate’) because V.K.Finn’s critique on the original name.

Hence the prominent problem of VKF method is how to encode training examples features by bitsets. Then the system simply concatenates the bitset representations of single features into bitset encoded whole object.

2.1. Continuous attributes

At first, VKF method applies the approach by analogue with J.R. Quinlan's technique in C4.5 decision tree algorithm [8].

Let $E = O \cup C$ be a disjoint union of training examples O and counter-examples C. Interval of values of continuous attribute $V: E \to \mathbb{R}$ generates three subsets

$E[a, b) = \{e \in E | a \leq V(e) < b\}$;

$O[a,b) = \{e \in O | a \leq V(e) < b\}$;
$C[a,b) = \{e \in C | a \leq V(e) < b\}$.

Definition 1. ***Entropy*** of interval of values of continuous attribute $V: E \to \mathbb{R}$ is

$$ent[a,b) = -log\left(\frac{|O[a,b)|}{|E[a,b)|}\right) \cdot \frac{|O[a,b)|}{|E[a,b)|} - log\left(\frac{|C[a,b)|}{|E[a,b)|}\right) \cdot \frac{|C[a,b)|}{|E[a,b)|}. \quad (1)$$

Mean information for partition of interval of values of continuous attribute $V: E \to \mathbb{R}$ is

$$inf[a,b) = ent[a,r) \cdot \frac{|E[a,r)|}{|E[a,b)|} + ent[r,b) \cdot \frac{|E[r,b)|}{|E[a,b)|} \quad (2)$$

Threshold is a value with minimal mean information.

For continuous attribute$V: E \to \mathbb{R}$denote $a = minV$ by v_0 and let v_{l+1} be an arbitrary number greater then $b = maxV$. Thresholds $\{v_1 < \cdots < v_l\}$ are computed sequentially by splitting the largest entropy subinterval.

Then VKF method uses $2l$ bits to encode the attribute value that falls between these thresholds.

Definition 2. For each $1 \leq i \leq l$ indicator (Boolean) variables δ_i^V, σ_i^V correspond to

$$\delta_i^V(e) = 1 \Leftrightarrow V(e) \geq v_i \quad (3)$$
$$\sigma_i^V(e) = 1 \Leftrightarrow V(e) < v_i \quad (4)$$

Then string $\delta_1^V(e) \ldots \delta_l^V(e) \sigma_1^V(e) \ldots \sigma_l^V(e)$ is a ***bitset representation*** of continuous attribute V on element $e \in E$.

It can be easily proved that the bit-wise multiplication of such bitsets encodes that a value falls into the convex union of corresponding intervals. Empty (all zeroes) bitset corresponds to the trivial fact that a value falls between $minV$ and $maxV$.

For additional information on VKF experiment with training objects described by continuous features, see [7].

2.2. Discrete attributes

For discrete attribute assume that its values form a finite semilattice. There is similar result for infinite case, however the additional property of completeness is needed. For computer applications finite case is sufficient, of course.

Addition of the top element, if it absents, transforms the semilattice into a lattice.

Definition 3. A subset $S \subseteq L$ of a lattice L is called $\vee$-***dense***, if any element has representation $x = \vee X$ for some subset $X \subseteq S$.

Element called $\vee$-***irreducible*** if and for any $y, z \in L$ $y < x$ and imply $y \vee z < x$.

It is easy to check

Lemma 1. ***Any superset of all -irreducible elements of a finite lattice forms*** $\vee$***-dense subset.***

Definition 3. Let $S = \{s_1, \ldots, s_k\} \subseteq L$ be -dense subset of finite lattice L. Indicator (Boolean) variables σ_i correspond to

$$\sigma_i^V(x) = 1 \Leftrightarrow x \geq s_i \quad (5)$$

Then string $\sigma_1(x) \ldots \sigma_k(x)$ is a ***bitset representation*** of discrete attribute of element

Proposition 2. ***Let be a*** $\vee$***-dense subset of finite lattice*** L***. Then the lattice of all candidates for sample described by bitsets isomorphic to*** L***.***

For the formal representation of the algorithm of encoding of discrete values attributes and proof of its correctness see [5, 6]. Some results of VKF experiments with discrete attributes training samples are described in [5].

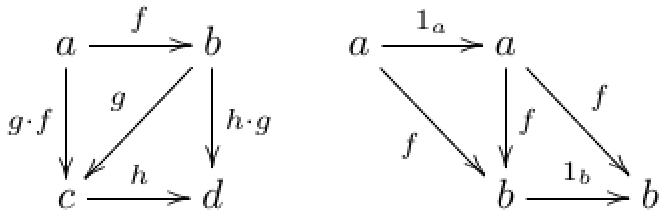
3. Category of semilattices

Local similarity is a binary operation on a set X that encompasses a set of objects, i.e. it is a mapping . Elements of the set X shall be called ‘fragments’.

For the result of similarity between several objects to be independent on their ordering, the similarity operation must satisfy the semilattice axioms: the laws of associativity, commutativity, and idempotency.

It is clear that in this case we can define a global similarity operation that for any subset $S \subseteq X$ generates the greatest lower bound $\wedge S \in X$. In other words, there is a mapping $\wedge: PX \to X$, where ***PX*** is the power-set (it equals the set of all subsets) for X.

We recall that a category ***C*** consists of a family of 'objects' and a family of 'arrows', where each object c has an arrow $1_c: c \to c$ and each pair of arrows $f: a \to b, g: b \to c \in C$ generates an arrow $g \cdot f: a \to c \in C$, where $f \cdot 1_a = f$, $1_b \cdot f = f$ and the composition is associative $h \cdot (g \cdot f) = (h \cdot g) \cdot f: a \to d$, if it is defined in ***C***, i.e., the following diagrams are commutative:



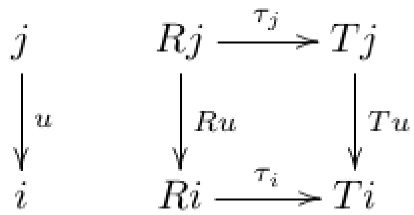
A diagram is 'commutative' if all compositions of arrows with common source and target objects are equal to each other.

We recall that a functor $T: J \to C$ from category ***J*** to a category ***C*** consists of a T objects function that assigns each object $j \in J$ to an object $Tj \in C$ and T arrow function that assigns each arrow $u: j \to i \in J$ to an arrow $Tu: Tj \to Ti \in C$, where $T(1_j) = 1_{Tj}$, $T(v \cdot u) = Tv \cdot Tu$ (when the composition $v \cdot u$ is defined in ***J***).

The category ***Set*** of sets and mappings between them allows endofunctor $P: Set \to Set$ that maps a set X to a power-set $PX = \{A \subseteq X\}$ and a mapping $f: X \to Y$ to $Pf: PX \to PY$, where $Pf(A) = \{f(x) | x \in A\} \subseteq Y$, for any $A \subseteq X$.

We will encounter below a 'forgetful] functor $F: Lat \to Set$, which maps a semilattice $\langle X, \wedge \rangle$ to its domain X and homomorphism of semilattices $f: \langle X, \wedge \rangle \to \langle Y, \wedge \rangle$ to mapping $Ff: X \to Y$.

We recall that a 'natural transformation' $\tau: R \to T$ from functor $R: J \to C$ to functor $T: J \to C$ is a function that assigns each object $j \in J$ to an arrow $\tau: Rj \to Tj \in C$ in such a way that for each arrow $u: j \to i \in J$ the following diagram is commutative:

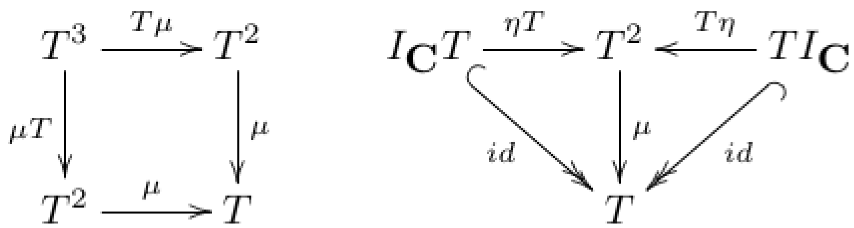


There exists a natural transformation η from an identity functor I_{Set} to a functor P with components $\eta_X: X \to PX$ that map each $x \in X$ to a single-element subset $\{x\} \in PX$.

There also exists a natural transformation $\cup: PP \to P$ with components $\cup_X: PPX \to PX$ that map each family $S \subseteq PX$ of subsets to their union$\cup \{A | A \in S\} \in PX$.

Each endofunctor $T: C \to C$ defines the compositions $T^2 = T \cdot T: C \to C$ and $T^3 = T^2 \cdot T: C \to C$. Let $\mu: T^2 \to T$ be a natural transformation with components $\mu_c: T^2 c \to Tc$ for each $c \in C$. Then $T\mu: T^3 \to T^2$ denotes a natural transformation with components $(T\mu)_c = T(\mu_c): T^3 c \to T^2 c$; the transformation $\mu T: T^3 \to T^2$ has components $(\mu T)_c = \mu_{Tc}: T^3 c \to T^2 c$.

We recall that a monad $\langle T, \eta, \mu \rangle$ in a category ***C*** consists of a functor $T: C \to C$ and two natural transformations $\eta: I_C \to T$ and $\mu: T^2 \to T$ that make the following diagrams commutative:



It is easy to check that the triple $\langle P, \eta, \cup \rangle$ defines the monad in the category ***Set***. The necessary identities $\cup \cdot (\cup P) = \cup \cdot (P \cup): P^3 \to P$ and $\cup \cdot (\eta P) = id = \cup \cdot (P\eta): P \to P$ correspond to the equality $\cup_{i \in \cup \{I_j | j \in J\}} S_i = \cup_{j \in J} \cup_{i \in I_j} S_i$ and the identities $\cup \{A | A \subseteq S\} = S = \cup \{\{x\} | x \in S\}$, respectively.

For the monad $\langle P, \eta, \cup \rangle$ in the category ***Set***, a category of algebras ***Lat*** consists a set of all pairs $\langle X, \wedge \rangle$, where the object (set) X is called 'domain' of the algebra and the morphism $\wedge: PX \to X$ is called 'structural mapping'; the identities

$$\wedge \cdot \cup_X = \wedge \cdot (P \wedge): P^2 X \to X \tag{6}$$

and

$$\wedge \cdot \eta_X = id_X: X \to X \tag{7}$$

must be true.

Lemma 3. ***The class of algebras*** $\langle X, \wedge \rangle$ ***over the monad*** $\langle P, \eta, \cup \rangle$ ***in the category Set coincides with complete semilattices.***

Proof. The structural mapping $\wedge: PX \to X$ defines the partial order $x \le y \Leftrightarrow \wedge \{x, y\} = x$. Antisymmetry: from $x \le y$ and $y \le x$ follows $x = \wedge \{x, y\} = \wedge \{y, x\} = y$. Reflexivity holds since identity (7): $\wedge \{x\} = \wedge \eta(x) = x$. Transitivity: from $x \le y$ and $y \le z$ follows $\wedge \{x, z\} = \wedge \{\wedge \{x, y\}, \wedge \{z\}\} = \wedge \{x, y, z\} = \cup \{\{x\} \cup \{y, z\}\} = \wedge \{\wedge \{x\}, \wedge \{y, z\}\} = \wedge \{x, y\} = x$.

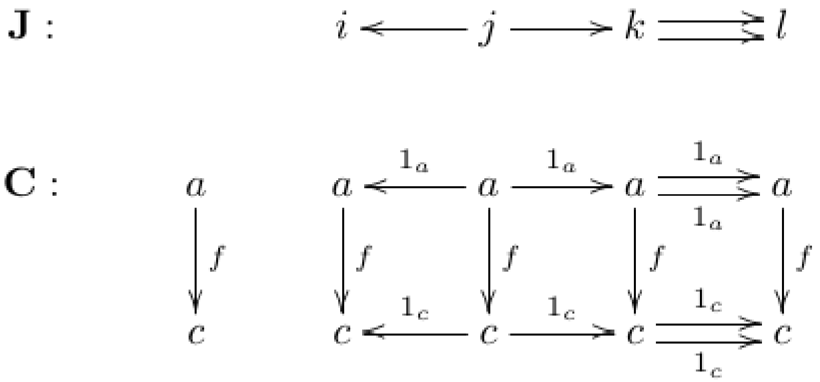
Let us prove that $\wedge S$ is the greatest lower bound for $S \subseteq X$. For each $x \in S$ it is true that $S = S \cup \{x\}$. From identity (6): $\wedge \cdot (\cup_X) = \wedge \cdot (P \wedge)$ follows $\wedge \{\wedge S, x\} = \wedge \{\wedge S, \wedge \{x\}\} = \wedge (S \cup \{x\}) = \wedge S$, which means $\wedge S \le x$. Let $\wedge \{y, x\} = y$ (i.e. $y \le x$) be true for each $x \in S$. Then $\wedge \{\wedge S, y\} = \wedge (S \cup \{y\}) = \wedge (\cup \{\{x, y\} | x \in S\}) = \wedge \{y\} = y$, i.e. $y \le \wedge S$.

The reverse statement that a complete semilattice $\langle X, \wedge \rangle$ is an algebra over the monad $\langle P, \eta, \cup \rangle$ is a simple exercise. ■

4. Cartesian products as limits of small functors

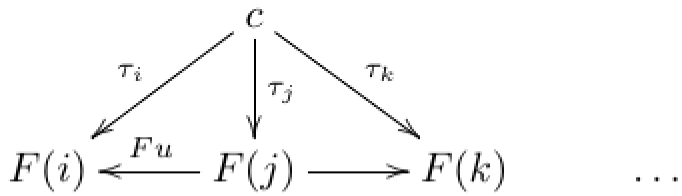
Let us denote the category of algebras over the monad $\langle P, \eta, \cup \rangle$ in the category ***Set*** as ***Lat***. This category is Cartesian-closed, i.e. it contains limits of small functors, whose definition is given below.

The 'diagonal' functor $\Delta: C \to C^J$ maps each object $c \in C$ to a constant functor $\Delta c: J \to C$, which on any object $j \in J$ takes the value c and on any arrow $u: j \to i \in J$ takes the value $1_c: c \to c \in C$. If $f: a \to c$ is some arrow from **C**, then Δf is a natural transformation $\Delta a \to \Delta c$, which on any object $j \in J$ takes the value $f: a \to c$.



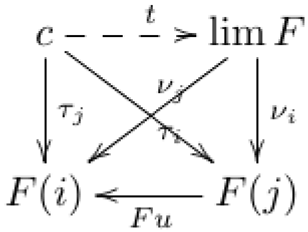
Let us call the natural transformation $\tau: \Delta c \to F$ from a constant functor Δc into some functor $F: J \to C$ a 'cone' with base F and vertex $c \in C$.

Since the values of the functor $\Delta c: J \to C$ can be reduced to $c \in C$, the natural transformation $\tau: \Delta c \to F$ for each object $j \in J$ consists of an arrow $\tau_j: c \to F(j)$ such that for any arrow $u: j \to i \in J$ the following diagram is commutative:



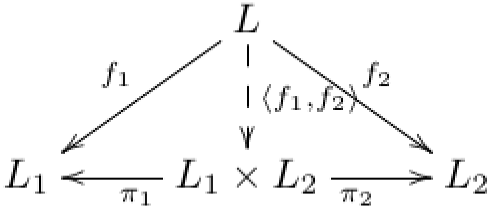
The limit of the functor $F: J \to C$ is a universal arrow $\langle r, \nu \rangle$ from Δ to F (it means the natural transformation $\nu: \Delta r \to F$). The object $r \in C$ is usually denoted as $\underset{\leftarrow}{lim}\, F$.

The natural transformation $\nu: \Delta r \to F$ is universal among natural transformations $\tau: \Delta c \to F$, where $c \in C$. In other words, the transformation $\nu: \Delta r \to F$ is a cone with base F and vertex $r \in C$ such that for any cone τ with base F and vertex c, there exists a unique arrow $t: c \to r$ such that $\tau_j = \nu_j \cdot t$ for all $j \in J$



An example of a limit is a 'Cartesian product', in which case the category $J = \langle \{1,2\}, \emptyset \rangle$ contains two objects and the empty set of arrows.

Then universal property of Cartesian product has the following form



In Lat limit $\underset{\leftarrow}{lim}\, F$ consists all 'threads', i. e. all elements of Cartesian product $\alpha \in \prod_{j \in J} L_j$ with coordinates α_j such that $Fu(\alpha_j) = \alpha_i$ for all arrows $u: j \to i \in J$. Operation $\wedge$ between threads is a restriction of Cartesian product one, i. e. it is computed component-wise.

The combination of lattices of attribute values can be described with a certain functor $F: J \to Lat$, where an object of J represents one of the attributes (generally speaking, arbitrarily complex), and an arrow $u: j \to i$ is a homomorphism of semilattices.

For instance, 'amalgam' of two lattices with respect to third one corresponds to the limit of functor from the category $J = \langle \{1,2,3\}, \{u: 1 \to 3, v: 2 \to 3\} \rangle$. This construction occurs when there exist two representation languages (for example, with different set of features) that can be reduced to common language (to the common subset of attributes). Then amalgams is useful to find a pairs of hypotheses in different languages that corresponds to same one in common description.

Another useful case is a limit of extending family of Cartesian products:

$$L_1 \xleftarrow{\pi_1^{1,2}} L_1 \times L_2 \xleftarrow{\pi_{1,2}^{1,2,3}} (L_1 \times L_2) \times L_3 \longleftarrow \dots$$

The limit of functor coincides with $\prod L_j$ together with component-wise similarity. If a lattice L_j represents a values of j-th attribute, the procedure corresponds to sequential variant of VKF method, when features appear one-by-one. It is possible even more general situation when expert obtains features in blocks. Then it needs to consider projection $\pi_U^V: \prod_{j \in V} L_j \to \prod_{j \in U} L_j$, where $V = U \cup \{J_i\}$, when J_i enumerates features in i-th block and U contains previously obtained attributes.

5. Free semilattices and polar operation on them

To define 'global similarity' (or, in the terminology of FCA [2], 'polar' operation), we need to establish the freeness of the algebra $\langle PS, \cup \rangle$ on a set S of training examples through conjugation of forgetful $F: Lat \to Set$ and generating $G: Set \to Lat$ functors.

We recall that a conjugation between categories C and B is a triple $\langle F, G, \varphi \rangle$, where $F: C \to B$ and $G: B \to C$ are functors, and the bijection φ assigns each arrow $h: Gc \to b$ to an arrow $\varphi h: c \to Fb$ conjugate to h on the right, and for all arrows $f: c' \to c$ and $g: b \to b'$ the conditions of naturalness are met:

$$\varphi(g \cdot h) = Fg \cdot \varphi h, \varphi(h \cdot Gf) = \varphi h \cdot f \quad (8)$$

This is equivalent to naturalness of the transformation φ^{-1}, that is, for all $f: c' \to c$, $g: b \to b'$, and $k: c \to Fb$ the following relations are true:

$$\varphi^{-1}(k \cdot f) = \varphi^{-1}k \cdot Gf, \varphi^{-1}(Fg \cdot k) = g \cdot \varphi^{-1}k \quad (9)$$

For any object $c \in C$, we consider $\eta_c: c \to FGc$ as an image of the arrow $1_{Gc}: Gc \to Gc$ at mapping φ. These arrows are components of a natural transformation η of an identity functor I_C to a functor $FG: C \to C$.

The bijection φ can be expressed in terms of arrows $\eta_c: c \to FGc$, namely, as

$$\varphi h = Fh \cdot \eta_c \quad (10)$$

for all $h: Gc \to b$.

Indeed, in view of the conditions of naturalness, $\varphi h = \varphi(h \cdot 1_{Gc}) = Fh \cdot \varphi 1_{Gc} = Fh \cdot \eta_c$.

Similarly, there exists a natural transformation ε of a functor $GF: B \to B$ into an indentity functor I_B as a set of arrows $\varepsilon_b: GFb \to b$: images of arrows 1_{Fb} at inverse bijection φ^{-1}.

Conversely, the bijection φ^{-1} can be expressed in terms of arrows $\varepsilon_b: GFb \to b$ by the formula

$$\varphi^{-1}k = \varepsilon_b \cdot Gk \quad (11)$$

for all $k: c \to Fb$.

Theorem 4. For the monad $\langle P, \eta, \cup \rangle$ in the category ***Set***, a conjugation $\langle F, G, \varphi \rangle$ exists, where $F: Lat \to Set$ is a forgetful functor $F\langle X, \wedge \rangle = X$, $G: Set \to Lat$ is a generating functor $GS = \langle PS, \cup \rangle$, $\eta: I_{Set} \to FG$ is a natural transformation with $\eta_X: X \to PX$defined as $\eta_X(x) = \{x\}$, and $\varepsilon: GF \to I_{Lat}$ is a natural transformation with $\varepsilon_{\langle L, \wedge \rangle} = \wedge: PL \to L$.

Proof. The functor$F: Lat \to Set$ forgets the structural mapping $F\langle X, \wedge \rangle = X$ of the semilattice.

For each set S the pair $\langle PS, \cup \rangle$ is a free complete semilattice over S by reason of the associative law and the presence of left and right units in the monad $\langle P, \eta, \cup \rangle$. Therefore, the correspondence $GS = \langle PS, \cup \rangle$ does in fact define the functor $G: Set \to Lat$, as stated.

Then $FGS = F\langle PS, \cup \rangle = PS$, so the unit $\eta_X: X \to PX$ of the monad $\langle P, \eta, \cup \rangle$ is a natural transformation $\eta: I_{Set} \to FG$.

On the other hand, $GF\langle X, \wedge \rangle = GX = \langle PX, \cup \rangle$. The identity (6): $\wedge \cdot (\cup_X) = \wedge \cdot (P \wedge)$ means that the structural mapping $\varepsilon_{\langle L, \wedge \rangle} = \wedge: PL \to L$ is a homomorphism of semilattices. This results in the natural transformation $\varepsilon: GF \to I_{Lat}$.

The remaining identities for conjugation have the forms and . The first one $\cup_X \cdot (P\eta_X) = id_{PX}: PX \to PX$ coincides with the equation $S = \cup \{\{x\} | x \in S\}$; the second is exactly the condition (7). ■

According to the proved theorem, the conjugation $\langle F, G, \varphi \rangle$ determines the isomorphism $\varphi^{-1}: Set(S, F\langle X, \wedge \rangle) \to Lat(\langle PS, \cup \rangle, \langle X, \wedge \rangle)$, for which the function of descriptions of training examples $f: S \to X$ in a domain of a lattice $\langle X, \wedge \rangle$ can be extended on all their subsets PS as shown in (11): $\varphi^{-1}(f) = \wedge \cdot Pf: \langle PS, \cup \rangle \to \langle X, \wedge \rangle$.Hence, each subset $A \subseteq S$, called a list of parents, maps to an element $\varphi^{-1}(f)A \in X$, called the polar or global similarity of the list of parents.

6. Conclusion

Although the original construction fits under the concept of a Ganter–Kuznetsov pattern structure [10], the categorical formalization replaces the concept of a projection between such structures with a homomorphism of algebras over the corresponding monad, which ensures invariance of the polar construction with respect to transformations of object names and semilattices of values.

Unfortunately, paper [10] contains a large number of incorrect statements (Proposition 1, Theorem 2). The invalidity of Proposition 1 from [10] was discovered by M.V. Samokhin [11]. The simplest counter-example was constructed by A.V. Buzmakov [12]. To eliminate this problem, the authors of [13] imposed various additional conditions on the definition of pattern structure projections, which are unnatural. The invalidity of the key Theorem 2 of [10] was established by T.B. Kaiser and S.E. Schmidt [14]; it remains invalid for the augmented definitions of projections from [13].

7. Acknowledgements

The author is grateful to L.A. Yakimova for helpful discussions and to his colleagues at the Dorodnicyn Computing Center of the Russian Academy of Science for their support. The critical comments and suggestions of Prof. S.O. Kuznetsov (HSE University, Moscow) were useful and the author would like to thank him for interest and discussions.

8. References

[1] V.K. Finn, Plausible Inferences and Plausible Reasoning, Journal of Soviet Mathematics, 56.1 (1991): 2201–2248.

[2] B. Ganter, R. Wille, Formal Concept Analysis. Springer, Berlin, 1999. doi:10.1007/978-3-642-59830-2.

[3] S.O. Kuznetsov, Mathematical Aspects of Concept Analysis, Journal of Mathematical Sciences 80.2(1996): 1654-1698.

[4] D.V. Vinogradov, Asymptotic probability of encountering an accidental similarity in the presence of counter examples, Automatic Documentation and Mathematical Linguistics. 51.1 (2017), pp. 42–44. doi:10.3103/S000510-5-517-01006-X

[5] D.V. Vinogradov, Machine Learning Based on Similarity Operation, in S.O. Kuznetsov et al. (Eds.) 16th Russian Conference RCAI 2018, volume 934 of Communications in Computer and Information Science, Springer Nature, Cham, 2018, pp. 46–59. doi:10.1007/978-3-030-00617-4_5

[6] D.V. Vinogradov, On Object Representation by Bit Strings for the VKF-Method, Automatic Documentation and Mathematical Linguistics. 52.3 (2018), pp. 113–116. doi:10.3103/S000510-5-518-03004-4

[7] D.V. Vinogradov, Continuous Attributes for FCA-based Machine Learning, in S.O. Kuznetsov et al. (Eds.) Proceedings of the 8th International Workshop "What can FCA do for Artificial Intelligence?" (FCA4AI 2020), volume 2729 of CEUR Workshop Proceedings, 2020, pp. 103–112.

[8] J.R. Quinlan, C4.5 Programs for Machine Learning, Morgan Kaufmann, San Francisco, 1993.

[9] S. MacLane, Categories for the Working Mathematician, Springer, NY, 1971.

[10] B. Ganter, S.O. Kuznetsov, Pattern structures and their projections, in H.S. Delugach, G. Stumme (Eds.) Proc. 9th Conference Conceptual Structures: Broadening the Base. ICSS 2001, volume 2120 of Lecture Notes in Computer Science, Springer, NY, 2001, pp. 129–142. doi: 10.1007/3-540-44583-8_10

[11] M.V. Samokhin, Machine learning on patterned structures, Ph.D. thesis (Eng.), VINITI, Moscow, 2006 (in Russian).

[12] A.V. Buzmakov, Modeling of processes with states of a complex structure based on lattices of closed descriptions, Ph.D. thesis (Eng.), HSE, Moscow, 2014 (in Russian).

[13] A.V. Buzmakov, S.O. Kuznetsov, A. Napoli, Revisiting pattern structure projections, in J. Baixeries, C. Sacarea, M. Ojeda-Aciego (Eds.) Formal Concepts Analysis Proc. ICFCA 2015, volume 9113 of Lecture Notes in Computer Science, Springer, 2015, pp. 200–215. doi: 10.1007/978-3-319-19545-2_13

[14] T.B. Kaiser, S.E. Schmidt, Some remarks on the relation between annotated ordered sets and pattern structures and their projections, in S.O. Kuznetsov et al. (eds) Pattern Recognition and Machine Intelligence. Proc. Conference PReMI 2011, volume 6744 of Lecture Notes in Computer Science, Springer, 2011, pp. 43–48. doi: 10.1007/978-3-642-21786-9_9