

Chapter 2

Information Data Structures

2.1 Overview

Data structures, orderings and applied classifications are generally defined on finite sets or sets of relations, which supposes that we know what are such entities (Sect. 2.3). But a part of classification research deals with data mining and the constitution of structured domains of concepts or objects. The fact that a mathematical structure, the Galois connection, contains a quasi-exhaustive information about the correspondence of two sets (Sect. 2.4) has suggested to use this structure in association with an order relation (Sect. 2.5) to initiate formal conceptual analysis (Sect. 2.6). But *formal* concepts are not *real* concepts. The exploration of concrete structures of objects has then led to the construction of formal (Sect. 2.7) and regional (Sect. 2.8) ontologies, using sometimes, as Barry Smith does, non-classical logics (the mereology of Lesniewski). In all this chapter, we study these models in relation to the main problems of classification and, finally, discuss (Sect. 2.9) the theories and results that have been introduced.

2.2 Historical Notes

Finite sets were first studied by Alfred Tarski (see [479]), a polish logician and one of the father founders of modern mathematical logic. From a set theoretic viewpoint, they have been investigated by the French mathematician Claude Frasnay (see [180, 181]) and it is now, obviously, the true domain of combinatorics (see [4]). Galois connection, the basic model for FCA (Formal Concept Analysis), has been rediscovered many times. Studied by Barbut and Monjardet (see [24]), and already used, with the Galois lattice, to structure philosophical data information in the 1980s by Parrochia (see [375]), it has already been anticipated in Salton's work (see [441]), where document/term lattices are essentially Galois lattices in the sense of FCA. Similarly, during the 1980s, Yuli Schreider (see [444]), from the Russian school of taxonomy (see [220]), seems to have developed a categorical formalization of the

Galois lattice. We can also mention that feature structure lattices, as used in linguistic Componential Analysis at the end of the 1990s, were very similar to concept lattices (see [146]). But as far as we know, the use of Galois lattices (now more often called “conceptual lattices”) in Data Analysis goes back to Barbut (see [22]), which proved also, in the same paper, that every lattice is Galois. “Formal concepts” have been popularized by Rudolf Wille (see [502]), but they were already known, under the name of “maximal rectangles”, by Kaufmann and Pichat (see [271]), and also by Parrochia (see [378]). Since that time, lattice theory and conceptual graphs (see [471]) have been used in many applied fields, from information science (see [403]) to technology design (see [384]), going through most of social sciences, medicine, biology, computer sciences, mathematics and so on.

Formal ontologies are generally explicit specifications of a conceptualization. However, as formal concepts are different from real concepts, the structure of the concepts of a real domain of objects may be very different from the structure of their formal concepts. This point necessitates constructions of explicit ontologies. Ontology began with the Greek philosophers (in particular, Aristotle), and has been developed in a quite philosophical manner from the metaphysics of Wolff (see [509]) to the *Logical Investigations* of Husserl (see [254]). After Frege, Russell and Wittgenstein, and some advances in mathematical logic, *formal* ontology became more and more a *formalized* ontology, which spreads out in many domains, and especially Computer Science and Artificial Intelligence with the birth of object oriented languages. Many authors like Sommers (see [467]), Grossmann (see [214]), Cocchiarella (see [93]), Poli or others might be quoted here. In applied domains, and particularly biology and medicine, regional ontologies have been constructed by Barry Smith (whose work will be studied below).

2.3 The Notion of a Finite Set

If you asked a non-mathematician what a *finite set* is, and how you can know that a set is *finite*, his answer would probably be that it is sufficient you enumerate the number of elements of the set, and the game is over. In a more convenient language, we get the classic following definition:

2.3.1 Some Definitions

Definition 2.3.1 A set E is said to be finite if there exists a bijection between E and a set $B = \{k \in \mathbb{N} : k < n\}$ for some $n \in \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers.

However, it has been observed for a long time that the above definition of “finite” only makes sense given the set of natural numbers, i.e. given an axiom of infinity. In fact, a foundational concern raises the problem: how to capture the concept of finiteness without making use of external objects like the natural numbers?

Of course one will say that concern soon became irrelevant, since a canonical way of representing natural numbers as sets has been devised. Nevertheless, quoting from [113], 155 “the definitions of finiteness remain a useful way of classifying non-well-orderable sets according to their cardinal-invariant properties”.

In the first part of the XXth century, the problem of the definition of finite sets was tackled by Tarski. In a paper, published in 1924, Tarski (see [479]) gives definitions of finite sets, some of which are still alive in more recent books (see [476], [178], 12–14). In fact, there are a lot of definitions in the literature, that Cruz (see [113], 156) has carefully collected. Expressed in modern language, we can find, in particular, the following ones:

Definition 2.3.2 A set X is:

1. I-finite if every non-empty family of subsets of X has a maximal (or minimal) element under inclusion.
2. Ia-finite if it is not the disjoint union of two non-I-finite sets.
3. II-finite if every family of subsets of X linearly ordered by inclusion contains a maximal element.
4. III-finite if there is no one-to-one map from $P(X)$ into a proper subset of $P(X)$.
5. IV-finite if there is no one-to-one map from X into a proper subset of X .
6. V-finite if $X = \emptyset$ or there is no one-to-one map from $X \times X$ into X .
7. VI-finite if X is empty or a singleton, or else there is no one-to-one map from $X \times X$ into X .
8. VII-finite if X is I-finite or it is not well orderable.

Tarski (see [479]) introduced definitions I, II, III and V, and definition VI has been attributed to him by Mostowski. Definitions Ia and VII are due to Lévy, while definition IV is due to Dedekind.

One of the benefits of Tarski’s approach is that the logician introduces (see [479], 33) the notion of *order* without resorting to the general notion of relation, and so, finite sets are defined by some general conditions of ordering on them.

Another definition is the one given by Howard and Yorke (see [249]):

Definition 2.3.3 A set X is D-finite if it is empty, a singleton, or if $X = X_1 \cup X_2$, where $|X_1|, |X_2| < |X|$ (such a set is also called decomposable).

Remark 2.3.4 Such a definition of finiteness seems implicitly assumed by Barry Smith (see below) when he introduces the notion of a “granulary partition”, which is, at first, a partition defined on a decomposable set.

Finally, Cruz (et al.), in a Tarski style, have introduced definitions of finiteness derived from classes of order relations:

Definition 2.3.5 A set X is:

1. P-finite if it contains a maximal element under every partial ordering.

2. L-finite if it contains a maximal element under every linear ordering.
3. W-finite if it contains a maximal element under every well ordering.
4. Tr-finite if it contains a maximal element under every tree ordering.

Next, they give a means of deriving new definitions of finiteness from old ones. Let Q stand for any of the notions given in this section.

Definition 2.3.6 A set X is:

1. Q^1 -finite if every set that X maps onto is Q -finite.
2. Q^2 -finite if every set that maps one-to-one into X is Q -finite.

Interesting cases are those where Q represents one of the notions of finiteness given in Definition 2.3.5.

Theorem 2.3.7 *If Q is any notion of finiteness, then a set X is Q^1 -finite if and only if every partition of X is Q -finite, and Q^2 -finite if and only if every subset of X is Q -finite.*

As a consequence, we can see that not only the existence of partial, linear or well ordered partitions, but the existence of a partition with a tree-structure (i.e. a hierarchy of classes) yields in fact a finite set. So finite sets are not only, as one says, finitely presentable objects of the category *Set* of all sets, or objects of the category *Finset* of all finite sets, which is, with the functions defined between them, essentially the subject matter of Combinatorics. Actually, we would be justified to assume the hypothesis that some implicit order is behind any idea of finiteness.

2.3.2 Implications Between Notions of Finiteness

Let us quote some implication relations among the previous definitions, which are closely connected. One proves easily the following theorems:

Theorem 2.3.8 (Lévy) $I \rightarrow Ia \rightarrow II \rightarrow III \rightarrow IV \rightarrow V \rightarrow VI \rightarrow VII$.

Theorem 2.3.9 (Truss) $III \rightarrow IV$ and $II \rightarrow IV$.

Theorem 2.3.10 (Howard and Yorke) $IV \rightarrow D \rightarrow VII$.

Theorem 2.3.11 (Cruz et al.) *If Q is any of the notions of finiteness defined above, then $Q^1 \rightarrow Q^2 \rightarrow Q$.*

Proof Given a set X , observe that every set that maps one-to-one into X can be mapped onto by X . So if every set that can be mapped onto by X is Q -finite (i.e. X is Q^1 -finite), then every set that maps one-to-one into X is Q -finite (i.e. X is

Q^2 -finite). $Q^2 \rightarrow Q$. Now let X be Q^2 -finite. It means that every set that maps one-to-one into X , including X itself, is Q -finite. $\square$

It appears that the most general definition of a finite set is Tarski's as well as any of Cruz's definition of a simple mapping on some ordering, whatever it is.

2.4 Galois Connection and the Whole Information

With respect to finite sets, we most of the time deal with implicit or explicit orders. So we have often to check the characteristics of those ordered sets and represent them by some appropriated data structures. We get there a first example of mapping from an ordered set into another one. This type of mappings are called "morphisms". Among the different kinds of morphisms studied, for example, by Caspard, Leclerc and Monjardet (see [80], 71)—whose names are: codings, embeddings, closures, spans, residual and residuated mappings...—one of them, the *Galois connection*, proves particularly useful in the field of classification and information science, as we shall see in the next sections.

Informally speaking, a Galois connection¹ is a relational structure, more precisely a particular correspondence between two partially ordered sets (or posets), two preordered sets, or even two classes. We shall only examine, in the following, the case of posets.

Definition 2.4.1 An order relation is a binary relation " $\leq$ " over a nonempty set P which is reflexive, antisymmetric, and transitive, i.e. for all a, b , and $c \in P$, we have:

1. $a \leq a$ (reflexivity);
2. if $a \leq b$ and $b \leq a$ then $a = b$ (antisymmetry);
3. if $a \leq b$ and $b \leq c$ then $a \leq c$ (transitivity).

Let us make some remarks:

1. If only reflexivity and transitivity hold, the relation $\leq$ is only a relation of pre-order (or quasi-order) on P .
2. For all a, b , distinct elements of $(P, \leq)$, if $a \leq b$ or $b \leq a$, then a and b are said to be "comparable". Otherwise they are "incomparable".

¹The notion of a Galois connection has its roots in Galois theory. By a fundamental theorem of Galois theory, there is a one-to-one correspondence between the intermediate fields between a field L and its subfield F (with appropriate conditions imposed on the extension L/F), and the subgroups of the Galois group $Gal(L/F)$ such that the bijection is inclusion-reversing:

$$Gal(L/F) \supseteq H \supseteq \langle e \rangle \quad \text{iff} \quad F \subseteq L^H \subseteq L, \quad \text{and} \\ F \subseteq K \subseteq L \quad \text{iff} \quad Gal(L/F) \supseteq Gal(L/K) \supseteq \langle e \rangle$$

3. If not all elements of $(P, \leq)$ are comparable, the relation $\leq$ is a *partial* order relation.

Definition 2.4.2 (Poset) A set with a partial order relation is called a partially ordered set (also called a poset).

Let us observe that:

1. If every two elements of a poset are comparable, the poset is called a totally ordered set or *chain*.
2. A poset in which every two distinct elements are incomparable is called an *antichain*.

Definition 2.4.3 Let $(P, \leq_P)$ and $(Q, \leq_Q)$ be two posets. A Galois connection between $(P, \leq_P)$ and $(Q, \leq_Q)$ is a pair of functions (f, g) with $f: P \rightarrow Q$ and $g: Q \rightarrow P$, such that, for all $p \in P$ and $q \in Q$, we have

$$f(p) \leq_Q q \quad \text{iff} \quad p \leq_P g(q).$$

One denotes a Galois connection between P and Q by $P \xrightarrow{f} Q$, or simply $P \multimap Q$.

If we define $\leq'_P$ on P by $a \leq'_P b$ iff $b \leq_P a$, and define $\leq'_Q$ on Q by $c \leq'_Q d$ iff $d \leq_Q c$, then $(P, \leq'_P)$ and $(Q, \leq'_Q)$ are posets, (the duals of $(P, \leq_P)$ and $(Q, \leq_Q)$). The existence of a Galois connection between $(P, \leq_P)$ and $(Q, \leq_Q)$ is the same as the existence of a Galois connection between $(Q, \leq'_Q)$ and $(P, \leq'_P)$. In short, one says that there is a Galois connection between P and Q if there is a Galois connection between two posets S and T , where P and Q are the underlying sets (of S and T respectively). With this, one may say quite clearly that a Galois connection exists between P and Q iff a Galois connection exists between Q and P .

We can make again some remarks:

1. Since $f(p) \leq_Q f(p)$ for all $p \in P$, then by Definition 2.4.3, $p \leq_P gf(p)$. Alternatively, we can write

$$1_P \leq_P gf \tag{2.1}$$

where 1_P stands for the identity map on P . Similarly, if 1_Q is the identity map on Q , then

$$fg \leq_Q 1_Q. \tag{2.2}$$

2. Suppose $a \leq_P b$. Since $b \leq_P gf(b)$ by the remark above, $a \leq_P gf(b)$ and so, by definition, $f(a) \leq_Q f(b)$. This shows that f is monotone. Likewise, g is also monotone.
3. Now back to inequality (2.1), $1_P \leq_P gf$ in the first remark. Applying the second remark, we obtain:

$$f \leq_Q fgf. \tag{2.3}$$

Next, according to inequality (2.2), $fg(q) \leq_Q q$ for any $q \in Q$, it is true, in particular, when $q = f(p)$. Therefore, we also have:

$$fgf \leq_Q f. \quad (2.4)$$

Putting inequalities (2.3) and (2.4) together, we get:

$$fgf = f. \quad (2.5)$$

Similarly,

$$gfg = g. \quad (2.6)$$

4. If (f, g) and (f, h) are Galois connections between $(P, \leq_P)$ and $(Q, \leq_Q)$, then $g = h$. To see this, observe that $p \leq_P g(q)$ iff $f(p) \leq_Q q$ iff $p \leq_P h(q)$, for any $p \in P$ and $q \in Q$. In particular, setting $p = g(q)$, we get $g(q) \leq_P h(q)$ since $g(q) \leq_P g(q)$. Similarly, $h(q) \leq_P g(q)$, and therefore $g = h$. By a similar argument, if (g, f) and (h, f) are Galois connections between $(P, \leq_P)$ and $(Q, \leq_Q)$, then $g = h$. Because of this uniqueness property, in a Galois connection $f = (f, g)$, f is called the upper adjoint of g and g the lower adjoint of f .

Remark 2.4.4 The pair of functions in the above Galois connection (as it is defined) are order preserving. One may also define a Galois connection as a pair of maps $f : P \rightarrow Q$ and $g : Q \rightarrow P$ such that $f(p) \leq_Q q$ iff $g(q) \leq_P p$, so that the pair f, g are order reversing. In any case, the two definitions are equivalent in that one may go from one definition to another, (simply exchange Q with Q^∂ , the dual of Q).

2.5 Galois Connection Associated to a Relation

One of the main characteristics of a Galois connection is to show the existence of dual substructures within two different ordered sets. When, moreover, these ordered sets are lattices, we get numerous additional properties. In the following, we only expose the case of the lattices of subsets of two sets.

2.5.1 Lattices and Morphisms of Lattices

Definition 2.5.1 Recall that a partial ordered set (or poset) $(T, \leq)$ is an inf-semilattice if every pair $\{x, y\}$ of elements of T has one and the same least upper bound (l. u. b.). Now, $(T, \leq)$ is a sup-semilattice if every pair $\{x, y\}$ of elements of T has one and the same greatest upper bound (g. u. b.). Then, $(T, \leq)$ is a lattice if it is both an inf-semilattice and a sup-semilattice. Note that the element 0_T is the minimum of T , and that the element 1_T is its maximum.

Definition 2.5.2 Let T and T' be two lattices, and f a mapping from T into T' . Then, the mapping f is:

1. a residuated mapping iff it is a sup-morphism satisfying $f(0_T) = 0_{T'}$;
2. a residual mapping iff it is an inf-morphism satisfying $f(1_T) = 1_{T'}$;
3. a Galois mapping iff $f(0_T) = 1_{T'}$ and for all $x, y \in T$, $f(x \vee x') = f(x) \wedge f(x')$.

Definition 2.5.3 Let P be a partial ordered set. A mapping ϕ is said to be extensive if $d_P \leq \phi$, i.e. if $x \leq \phi(x)$ for all $x \in P$.

Definition 2.5.4 Let P, Q be partially ordered sets. A mapping ϕ from P into Q , such that for all $x, x' \in P$, $x \leq x' \implies \phi(x) \geq \phi(x')$ is said to be antitone (the dual concept of an antitone mapping is an isotone mapping).

Definition 2.5.5 Let P, Q be partially ordered sets, and f a mapping from P into Q . f is a Galois mapping iff f is antitone, and if there exists an antitone mapping g from Q into P such that the mappings $\phi = gf$ and $\psi = fg$ are extensive.

2.5.2 Galois Lattice (Conceptual Lattice)

Let now E and F be two finite sets, $\underline{2}^E$ (respectively $\underline{2}^F$), the lattice of all the subsets of E (respectively F). Let also R be a relation such that $R \subseteq E \times F$. Let us set: for all $e \in E$, $f \in F$, $eR = \{f \in F : eRf\}$ and $Rf = \{e \in E : eRf\}$. Then, for all $A \subseteq E$, $B \subseteq F$:

$$f_R(A) = \{b \in F : aRb \text{ for all } a \in A\} = \bigcap \{eR : e \in A\},$$

$$g_R(B) = \{a \in E : aRb \text{ for all } b \in B\} = \bigcap \{Rf : f \in B\}.$$

One can verify that the pair (f_R, g_R) are order reversing and that we always have $A \subseteq g_R f_R(A)$ and $B \subseteq f_R g_R(B)$.

Let us set now $\phi_R = g_R f_R$ and $\psi_R = f_R g_R$.

As f_R and g_R are antitone and as the mappings ϕ and ψ are extensive, the pair (f_R, g_R) satisfies (2.4.5) and so, is a Galois connection between the lattices, $\underline{2}^E$ and $\underline{2}^F$.

$\phi_R = g_R f_R$ and $\psi_R = f_R g_R$ are the closures (of, respectively, $\underline{2}^E$ and $\underline{2}^F$), associated with this connection. In general, we shall write f, g, ϕ, ψ instead of f_R, g_R, ϕ_R, ψ_R .

We shall not prove the following theorem (see [80], 91):

Theorem 2.5.6 Let $R \subseteq E \times E'$, a relation between two sets E and E' , F a subset of E and H a subset of E' . With the previous notations, the following four conditions are equivalent:

1. F is a closed set of ϕ and $H = f(F)$;
2. H is a closed set of ψ and $F = g(H)$;
3. $H = f(F)$ and $F = g(H)$;
4. $F \times H \subseteq R$, F and H being maximal for this property.

Let us now give some definitions, which we shall find again later in the “formal concept analysis” of Rudolf Wille.

Definition 2.5.7 (Formal concept) A (formal) concept is a set $F \times H$ satisfying the conditions of Theorem 2.5.1.

Definition 2.5.8 (Formal context) When E is a set of (formal) concepts, and E' a set of attributes, the triple (R, E, E') such that $R \subseteq E \times E'$ is called a *formal context* of these concepts.

Remark 2.5.9 Caspard, Leclerc and Monjardet note $Gal(E, E', R)$ the set of concepts $F \times H$ of R .

Because of the properties of the Galois connection, if $F \times H$ and $F' \times H'$ are two concepts, we have $F \subseteq F'$ iff $H' \subseteq H$. The set $Gal(E, E', R)$ is then ordered by the relation:

$$\begin{aligned} F \times H \leq F' \times H' &\iff F \subseteq F' \\ &\iff H' \subseteq H. \end{aligned}$$

Definition 2.5.10 The ordered set $(Gal(E, E', R), \leq)$ is called the *Galois lattice* (or the *concept lattice*) of R .

This set is a lattice since, by definition, it is isomorphic to $\mathcal{F} = \phi(2^E)$ and dual of $\mathcal{G} = \phi(2^{E'})$. One can deduce its operations are given by:

$$(F \times H) \wedge (F' \times H') = (F \cap F') \times (\psi(H \cup H'))$$

and

$$(F \times H) \vee (F' \times H') = (\phi(F \cup F')) \times (H \cap H').$$

Example Table 2.1 (see [217], [80], 92) gives a binary relation R from a set E to a set E' , with $E = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and $E' = \{A, B, C, D, E, F, G\}$. Table 2.2 shows the corresponding Galois lattice with 15 elements. A set of animals is described by a set of attributes which are binary properties. A closed set E_i forms the class of all objects having in common all the properties of the closed set $E'_i = f(E_i)$, which is the *extent* of the concept $(E_i, E'_i) = E_i \times E'_i$. These classes form the Moore family $\mathcal{F}$. On the other side, E'_i is the set of all the properties that the elements of E_i have in common, i.e. the *intent* of the concept (E_i, E'_i) . These descriptions form the dual Moore family $\mathcal{G}$. So the Galois lattice (or concept lattice) of the relation R is the ordered set of all concepts generated by the relation R (see [80], 93).

This gives the matrix of Table 2.2.

And we get, finally, the Galois lattice of the relation R (Fig. 2.1).

Table 2.1 Table of the relation R

1	Ostrich	A	Laying eggs
2	Canary	B	Feathered
3	Duck	C	Having scales
4	Shark	D	With a nude skin
5	Salmon	E	Having teeth
6	Frog	F	Flying
7	Crocodile	G	Swimming
8	Barracuda	H	Air breathing

Table 2.2 Matrix of the relation R

	A	B	C	D	E	F	G	H
1	x	x						x
2	x	x				x		x
3	x	x				x	x	x
4	x			x	x		x	
5	x		x				x	
6	x			x			x	x
7	x			x	x		x	x
8	x		x		x		x	

2.6 On Formal Concept Analysis

Using the properties of previous formalism (Galois connection and Galois lattice) known, as we have seen, for a long time, Formal Concept Analysis (FCA) was introduced by Rudolf Wille (see [502]) and the group of Darmstadt (Germany), in the beginning of the 1980s. The project aimed to develop systematically a framework for the application of lattice theory in different fields. This method, now mainly used for the analysis of data, i.e. essentially for investigating and processing information, actually applies in many different realms like psychology, sociology, anthropology, medicine, biology, linguistics, computer sciences, mathematics and industrial engineering. Since the end of the 1990s, different monographs have been published (see [79, 185, 186]) to expose the mathematical foundations of the method, a part of which has been integrated now in basic textbooks on lattice theory (see [206], 591–605).

Wille started from the hypothesis that data is generally structured into units that are formal abstractions of concepts of human thought, allowing for meaningful and comprehensible interpretations. The adjective “formal” is precisely used to emphasize that these formal concepts are in fact mathematical entities, quite different from the concepts of human thought. From a philosophical point of view, already emphasized by the old Port-Royal logic (see [11, 140]), a concept is a unit of thoughts consisting of two parts, the *extent* and the *intent*. As is well known, the extent covers

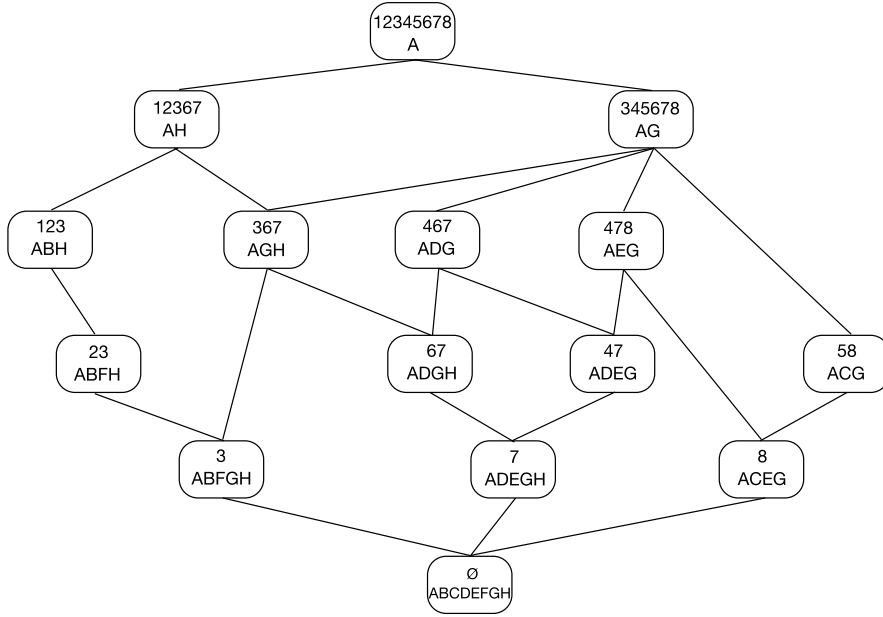


Fig. 2.1 Galois lattice of the relation R

all of the field of objects to which the concept applies, and the intent comprises all attributes valid for those objects. Hence, objects and attributes play a prominent role, together with some important relations like the hierarchical “subconcept-superconcept” relation between concepts, the implication between attributes, and the incidence relation saying that “an object has an attribute”. On this basis, Wille developed then his main ideas.

Remark 2.6.1 From now, the notion of “formal context” (which we have previously identified with the triple (R, E, E')) will be noted by the triple (G, M, I) where G is the set of objects, M the set of attributes, and I the relation R mentioned in Theorem 2.5.1. The notion of “formal concept”, previously noted $(F \times H)$, will be, from now, the pair (G_i, M_i) . The ordered set of all formal concepts previously noted $Gal(E, E', R)$, will be, from now, and according to the notation of Wille, the “concept lattice” of (G, M, I) .

As we can understand, the *formal context* $I \subseteq G \times M$ consists in fact of a set of objects G , a set of attributes M and an indication of which objects have which attributes.

Similarly, a *formal concept*, viewed as a pair (G_i, M_i) , is such that:

1. $G_i \subseteq G$;
2. $M_i \subseteq M$;
3. For every object in G that is not in G_i , there is an attribute in M_i that the object does not have;

4. For every attribute in M that is not in M_i , there is an object in G_i that does not have that attribute.

As we have seen, a context may be described as a table, with the objects corresponding to the rows of the table, the attributes corresponding to the columns of the table, and a Boolean value (in the example represented graphically as a cross) in cell (x, y) whenever object x has value y .

A concept, in this representation, forms a *maximal subarray* (not necessarily contiguous) such that all cells within the subarray are checked.

The fact that the concepts (G_i, M_i) defined above can be partially ordered by inclusion, such that their meet and join operations satisfy the axioms defining a lattice, is very important. By considering infinite meets and joins, analogously to the binary meets and joins defined in Definition 2.5.10, one sees that this is a *complete* lattice.² Conversely, any finite lattice may be generated as the concept lattice for some context. For, let L be a finite lattice, and form a context in which the objects and the attributes both correspond to elements of L . In this context, let object x have attribute y exactly when x and y are ordered as $x \leq y$ in the lattice. Then, the concept lattice (or Galois lattice) of this context is isomorphic to L itself. Such a construction may be interpreted as forming the Dedekind-Mac Neille completion of L , and is known to produce an isomorphic lattice from any finite lattice. In other words:

Definition 2.6.2 For every ordered set $P = (X, \leq)$, the Dedekind-Mac Neille completion of P is the Galois lattice $Gal(X, X, \leq)$ that we may denote as $Gal(P)$.

2.6.1 Formal Extensions

To be used as a tool for knowledge representation or knowledge discovery, for querying enlargement or refinement of sets of documents in information science, for building new semantic relations or a contextual logic for artificial intelligence, a concept lattice must allow us to go from an element to another, by the means of some operation, which is a kind of negation. However, modeling negation in a formal context is somewhat problematic.

Second, as the complement of a concept is not necessarily a concept, a group of attributes does not necessarily form a consistent entity. Some attributes may be irrelevant, the relationship to some object is unknown, or the object has only sometimes, usually or even never, this kind of attributes (see [72]). To solve the problem, Wille introduced what he has called “protoconcepts” or “semi-concepts” (see [503]) which are mathematical structures similar to formal concepts but whose mathematical properties are more complex and more difficult to describe. For the moment, the study of those structures is on-going.

²Recall that a complete lattice is a lattice in which every subset admits a greatest lower bound and a least upper bound. The lattice itself being one of its own subsets, it follows that it has a (unique) infimum and a (unique) supremum.

Concept Algebra of a Formal Context Nevertheless, since the concept lattice is complete, one can consider the join $(G_i, M_i)^\Delta$ of all concepts (G_j, M_j) that satisfy $G_j \subseteq G \setminus G_i$; or dually the meet $(G_i, M_i)^\nabla$ of all concepts satisfying $M_j \subseteq M \setminus M_i$. These two operations are known as weak negation and weak opposition, respectively.

This can be expressed in terms of the derivative functions. The derivative of a set $G_i \subseteq G$ of objects is the set $G'_i \subseteq M$ of all attributes that hold for all objects in G_i . The derivative of a set $M_i \subseteq M$ of attributes is the set $M'_i \subseteq G$ of all objects that have all attributes in M_i . A pair (G_i, M_i) is a concept if and only if $G'_i = M_i$ and $M'_i = G_i$. Using this function, weak negation can be written as

$$(G_i, M_i)^\Delta = ((G \setminus M'_i)'' , (G \setminus M'_i)')$$

and weak opposition can be written as:

$$(G_i, M_i)^\nabla = ((M \setminus G'_i)' , (M \setminus G'_i)'').$$

Concept lattice equipped with the two additional operations Δ and ∇ is known as the concept algebra of a context. Concept algebras are a generalization of power sets.

Weak negation on a concept lattice L is a *weak complementation*, i.e. an order-reversing map $\Delta : L \longrightarrow L$ which satisfies the following axioms:

1. $x^{\Delta\Delta} \leq x$;
2. $x \leq y \implies x^\Delta \geq y^\Delta$;
3. $(x \wedge y) \vee (x \wedge y)^\Delta = x$.

Weak composition is a dual weak complementation, which satisfies:

1. $x^{\nabla\nabla} \geq x$;
2. $x \leq y \implies x^\nabla \geq y^\nabla$;
3. $(x \vee y) \wedge (x \vee y)^\nabla = x$.

A (bounded) lattice such as a concept algebra, which is equipped with a weak complementation and a dual weak complementation, is called a *weakly dicomplemented lattice*. Finite distributive weakly dicomplemented lattices are isomorphic to concept algebras (see [288], 77).

Extensions to Many-Valued Contexts As we can guess, the basic data type of a formal context, including a set of attributes which are binary properties, is not the one occurring most frequently in the applications of FCA. As Ganter and Wille remark (see [206], 596), very often, data is recorded in the form of a tabular or relational data base described by the following definition:

Definition 2.6.3 A many-valued context (G, M, W, I) consists of a set G (of “objects”), a set M (of “many-valued attributes”), a set W (of “attributes values”) and a ternary relation $I \subseteq G \times M \times W$ satisfying:

$$(g, m, v) \in I, \quad (g, m, w) \in I \implies v = w.$$

In this case, the method consists of transforming this many-valued context into a “one-valued” one, i.e. into a formal context in the above sense, for constructing its concept lattice. The transformation procedure, called *conceptual scaling*, is, of course, not unique, and contains numerous degrees of freedom that can be interpreted in different ways. The simplest version of it is what Wille called “plain scaling”. It consists, for each value m of an m -valued context (G, M, W, I) , of constructing the scale $S_m = (G_m, M_m, W_m, I_m)$, the *derived context* of (G, M, W, I) being given as (G, N, J) , where N is just the Cartesian product $\{m\} \times M_m$, for each m , the relation J being each time particularized.

Fuzzy Concept Lattices and Rough Sets Other generalizations deal with Pawlak’s rough sets and fuzziness. Research in the use of concept lattices for knowledge discovery and data mining has shown that classic concept lattices are often too rigid structures to apply in real domains. Various approaches have been proposed to create fuzzy formal contexts and transform these ones into fuzzy concept lattices. Several approaches exist, for example, to create fuzzy concept lattices. One of these, demonstrated with bioinformatics data, specifically using gene annotation data files, has been developed in a recent paper (see [112]). The evidence code specified with an annotation is translated into a numeric value in $(0, 1)$ and is interpreted as the degree of association between the gene or gene product and the annotating Gene Ontology term. These degrees of association are used to create the fuzzy formal context which can then be used to create a fuzzy concept lattice.

Another way to reduce the gap between FCA and applications is to use another concept of fuzziness, which is the one we find in “Rough set theory” (see [389, 390]). Instead of formal concepts and concept lattices, which are the basic tools of FCA, the central notions of Rough set theory are the indiscernibility of objects with respect to a set of properties and the induced approximation operators. Different proposals have been made to combine the two theories in a common framework. On the one hand, one can introduce the notion of concept lattice into Rough set theory. On the other hand, one can introduce the notion of approximation operators into Formal Concept Analysis. This last viewpoint may be processed on the ground that one formulation of Rough set theory can be developed based on a binary relation between two universes, a finite set of objects and a finite set of properties. Rough set approximation operators (introducing modalities of necessity or possibility for the membership of elements in a set) are then defined with respect to the binary relation and several concept lattices may be constructed, based on approximation operators. Though different from the classic concept lattice, they are related to it. Through the study of these new lattices, one can then obtain an in-depth understanding of a conceptual data analysis (see [514]).

2.6.2 Applied Fields

Ganter and Wille, in different talks, have listed the many tasks that can be performed with their approach: exploring, searching, recognizing, identifying, some of these

operations being also important in classification. Similarly, the Darmstadt group has worked on selected projects in different disciplines, that illustrate the multitude of possible applications of FCA:

“*Medicine*: diabetes of children; *psychosomatics*: repertory grids of anorexic patients, *psychology*: children’s concept development, *musicology*: audiovisual perception of music, *politics*: international cooperations, *linguistics*: semantic structures of lexical databases, *information science*: retrieval system for a library, *computer science*: software reengineering, *electronics*: improving chip production, *civil engineering*: retrieval systems concerning laws and regulations, *ecology*: water pollution, *biology*: color perception.” (See [206], 597.)

Let us talk essentially of information science and, more precisely, of classification theory, which is the point.

FCA and Information Retrieval There was for a long time a theoretical interest for lattices in information retrieval. As we shall see (Chap. 3), Fairthorne, Hillman, Salton and others have tried to develop such an approach. But, as Priss said, none of these works resulted in practical implementations and, for a long time, the dominant mathematical model of information retrieval was the vector space. Interest for lattices was again spurred by Godin (see [192]), who concluded that the performance between Boolean queries and lattice navigation was similar and both better than the use of hierarchical classification. Then, Carpineto and Romano (see [79]) showed that FCA can serve three purposes in information retrieval: first, query refinement (and lattice structures can be used there to make suggestions); second, integration of querying and navigation (or browsing): users can navigate from a node to related ones. Third, integration of thesaurus hierarchy and concept lattice, an idea discussed by different researchers but probably yet to be perfect. In the beginning of the 2000s, some FCA software (like Credo or Mail-Sleuth) appeared to show a promise for applications in information retrieval. However, as Priss has shown, though Credo has been applied to thousands of documents in a small library (see [424]), in fact, it is not suited for direct manipulations of very large data sources. We can add that, as Milli (and others) have also shown, in general, controlled vocabularies (as those already used in a faceted classification) are most of the time too restrictive. Therefore, the hope that concept lattices could be of great use in library organization remains a dream. As distributive lattices are also concept algebras, the approach of Wille does not seem to go far beyond the research of Hillman (see below, Chap. 3).

Taxonomy The application of concept lattices to taxonomy began to develop in the Russian school before Wille popularized them through the so-called “formal concept analysis”.

Yuli Schreider (see [444]), influenced by Russian biologist and philosopher A.A. Liubishchev, realized that there was a deep parallelism between systematics and semiotics, the theory of signs. Description of a new “taxon” in taxonomy is very similar to the development of a new notion in language. Both are based on

similarity relationships among objects. This idea turned the interests of Schreider toward the mathematical theory of relations and models. A second largest intellectual stimulus came from his closest friend, Sergei Meyen, who developed a hierarchical approach to taxonomic similarity. By comparing organisms, one simultaneously compares their homological parts, which Meyen called “merons” (from the Greek *meros*, part).

The method for partitioning internal systems and classifying their homological parts (or “merons”) was named by Meyen “Meronomy” (see [343]). Meronomy was supposed to help the construction of natural classes. A taxonomy is called *natural* if it allows one to predict essential characteristics of objects from the fact that they belong to a specific taxon. Natural classes of whole systems form a taxon, and a class of parts of systems (these parts may belong to various systems or to the same system) is called a “meron”.

For Schreider and Meyen (see [344]), the relationship between taxonomy and meronomy is that each taxon can be characterized by some kind of “archetype”, true for all individuals (or species) that belong to this taxon. Such an assumption is equivalent to the existence of a “concept” for each notion in the language.

Indeed, Schreider considers an archetype as a frame built on the set of merons. Homology between parts of systems is the consequence of sharing the same archetype. Each part of a system corresponds to some meron in the archetype. Parts in different systems that correspond to the same meron in the archetype are homologous.

Both taxons and archetypes have a partial order. Taxons are ordered by inclusion. For example, the taxon of chimps is included in the taxon of primates. On the other side, Archetype A is greater than archetype B ($A > B$) if all merons and relations between them in B are also present in A .

One can prove that there exist antitone mappings f and g between the poset of taxons and the poset of archetypes, and that fg and gf are extensive, so that there is a Galois connection between the posets (Schreider has described it, in fact, in terms of Category theory).

But a direct application of concept lattices to taxonomy was performed in 1997 with Barwise and Seligman’s information flow model (see [28]), which incorporates relational structures and an explicit representation of context. This model presents information flow as *infomorphisms* between classifications, which are described in the following manner.

Definition 2.6.4 (Classification as a formal context) Formally, a classification (A, B, R) consists of a set A of *tokens*, a set B of *types* and a relation R between tokens and types (such that $R \subseteq A \times B$). For example, tokens can be objects and types can be attributes describing the objects.

As we can see, this definition is quite equivalent to a formal context in Formal Concept Analysis (see [185]).

So, a concept lattice can be constructed from a classification, as defined above.

Definition 2.6.5 (Concept lattice of a classification) Let $R(A)$ denote the set of tokens that all objects in A share and, vice versa, $R(B)$ denote the set of types that all tokens in B share. A concept is then defined as a pair (A_1, B_1) with $A_1 \subseteq A$ and $B_1 \subseteq B$, knowing that $R(A_1) = B_1$ and $R(B_1) = A_1$. The searched concept lattice is the set of concepts of the formal context (or classification) ordered by a subconcept-superconcept relation defined as:

$$(A_1, B_1) \leq (A_2, B_2) : \Longleftrightarrow A_1 \subseteq A_2 \text{ or equivalently } B_2 \subseteq B_1.$$

While concept lattices describe the relations within one classification, Barwise and Seligman's infomorphisms can be used to describe the information flow between different classifications.

Definition 2.6.6 (Infomorphism between classifications) An infomorphism is defined as a pair (f, f') of mappings between two classifications, (A_1, B_1, R_1) and (A_2, B_2, R_2) , such that for all tokens a of A_1 and for all types b of B_2 :

$$f(a)R_2b \Longleftrightarrow aR_1f'(b).$$

Remark 2.6.7 The two classifications are quasi-dual to each other, but not completely, since the tokens are mapped from the first one into the second one, whereas the types are mapped from the second one into the first one. The first one, whose tokens are mapped (in this case (A_1, B_1, R_1)), is called a *channel*. In this way, an *information flow channel* is itself a classification that contains some types and tokens from the other classifications.

Thus, the Barwise-Seligman model extends older Shannon-Weaver type models of information transmission, which are no longer passive message conductive cables. As Old and Priss have shown (see [368]), in Barwise-Seligman's model, the channel itself participates in the translation process between sender and receiver, which makes of it a "conceptual cable" between the two ends, the advantage being that the translation process itself can be separately analyzed and represented.

2.6.3 Some Debates

The theory of concept lattices is now a part of lattice theory, which yields some new theorems or extensions recently exposed by Ganter and Wille in a basic textbook (see [206], 591–605). In this book, after they have recalled the definitions of formal context and concept lattices, the authors introduce some basic theorems on concept lattices (for example, the fact that a concept lattice is complete), they explain how to find the smallest closed set larger than a given subset of objects, and prove some other theorems about sublattices and quotient lattices, subdirect products and tensor products, or lattice properties. One of the most interesting discussion comes from applications. In a clear-sighted way, Ganter and Wille point out one of the main problems of their method.

“Most other techniques of data analysis have as their aim to drastically reduce the amount of information given and to obtain a few “significant parameters”. By contrast, a concept lattice does not reduce complexity since it contains all the details of the data represented by the formal context. It is usually smaller than the power sets of G and M , since two closure systems of extents and intents contain only those sets that are, in this specific sense, meaningful. Nevertheless, it may be exponential in size, compared to the formal context. Complexity, therefore, is a problem, even though there are efficient algorithms and advanced application programs” ([206], 596).

This seems to generate some paradoxes: while FCA makes a big use of the Galois lattice, which is the basis of all constructions, the fact that this one contains too much information leads very often to give up a part of it (particularly when the aim is to get a clear classification). This means one has to truncate the concept lattice in some place (see, for example, [193]) in order to obtain a tree, which is generally done with the help of more or less convincing methods.

Another way to get out of trouble in very large databases is to limit the concept lattice to the most frequent concepts. This is the solution presented at the beginning of the 2000s with the concept of “iceberg concept lattice” (see [475]). The running example of this paper is the “mushroom” database.³ It consists of a database with 8416 objects (mushrooms) and 22 (nominally valued) attributes. A formal context is obtained by creating one (Boolean) attribute for each of the 80 possible values of the 22 database attributes. One obtains finally a data matrix with 8416 objects and 80 attributes, and an unreasonable concept lattice of the whole database with 32086 concepts. In order to simplify this one, the “iceberg concept lattice” will consist only of the top-most concepts of the concept lattice. These are the concepts which provide the most global structuring of the domain. The formal definition consists of fixing the size of what the author called a “minimum support” and then to make it decreasing to obtain more information. Because the support function is monotonously decreasing, the iceberg concept lattice is an order filter of the whole concept lattice, and thus, in general only a sup-semilattice. But when one adds a new bottom element, it becomes a lattice again. This makes it possible to apply the same algorithm for computing concept lattices and iceberg concept lattices. More recently, a generalization of iceberg concept lattices, the “Alpha Galois lattices”, seems to give better results (see [391]).

Other attempts have been developed for reducing the size of concept lattices, as, for example, fixing distances between formal concepts and introducing some additional clustering method (see [521]), or, in another style, extracting the most abstract concepts of a Galois lattice (see [297]).

But whatever the method used to get rid of the information overflow, the problem remains the same: why must we make use of an exhaustive structure, if it yields a so uncontrollable situation that we cannot use the whole of it, and are reduced, in the end, to some conventional mode of clustering?

³This database comes from the UCI KDD Archive (<http://kdd.ics.uci.edu/>).

In the same way, comparing different concept lattices representing results of experiment by the means of some distance or metric (whatever it is) (see [87]) is also quite problematic.

Finally, it seems that FCA may be especially useful in domains like intelligence help systems, or machine learning (see [311]), because the lattice operations allow the construction of new concepts by combining a few of the existing ones, taken as “primitive” (though this could be done, also, with the help of free distributive lattices, that are—as a recent theorem proves—completely equivalent to concept algebras). But the main classification problems remain unchanged. In particular, FCA cannot indicate what is the most relevant information in a complex field, and it cannot structure it in an economic way for the human mind.

Algorithms and Software Applications Whatever the final opinion we may have about FCA, we must recognize that the method has led to a number of algorithms. In 2001, Kuznetsov and Obiedkov (see [287]) had already surveyed a lot of algorithms that have been developed for constructing concept lattices. These algorithms vary in many details, but are in general based on the idea that each edge of the Hasse diagram of the concept lattice connects some concept C to the concept formed by the join of C with a single object. Thus, one can build up the concept lattice one concept at a time, by finding the neighbors in the Hasse diagram of known concepts, starting from the concept with an empty set of objects. The amount of time spent to traverse the entire concept lattice in this way is polynomial in the number of input objects and attributes per generated concept. New algorithms may be added now to the previous ones, some of them being even faster (see [88]).

We must observe also that many FCA software applications are now available. The main purpose of these tools is to generate the concept Lattice of a given formal context and the corresponding association rules. These tools are today academic and still under active development. Most of them are Java-based open-source applications like ConExp, Toscana, Coron or Lattice Miner. This last tool, one of the most famous, allows the generation of formal concepts and association rules as well as the transformation of formal contexts via apposition, subposition, reduction and object/attribute generalization, and it performs also the manipulation of concept lattices via approximation, projection and selection. Lattice Miner allows also the drawing of nested line diagrams, fulfilling so the main issue in the development of FCA tools, which is to visualize large concept lattices and provide efficient mechanisms to highlight patterns (e.g., concepts, associations) that could be relevant to the user.

2.7 Formal Ontologies

In this section, we shall first take a glance at the historical development of ontology, then we shall explain briefly what is a formal ontology, and finally, we shall explain the connection existing between formal ontologies and classifications.

From Classical to Formal Ontology Without going into details, let us say a few words about the evolution of the term “ontology” in the course of time. Under the name of “first philosophy”, ontology begins with Aristotle, but the word itself becomes of common use in philosophy during the Classical Age, and was only clearly defined by Christian Wolff (1679–1754). For this philosopher, ontology (from the Greek *ôs-ontos*, being) is the science or study of being, more specifically, a branch of metaphysics relating to the nature and relations of being, whose principles, as those of psychology, must precede the rules of logic (see [509], 17). However, this science “of something and nothing”, as Leibniz said in his *Introductio ad Encyclopaediam Arcanam* (see [107], 512), rapidly began to appear quite problematic, since most situations revealed that it was very easy to confuse imaginations with real things.

So, Emmanuel Kant drastically reduced the aim of ontology. For him, “ontology is that science (as part of metaphysics), which consists of a system of all concepts of the understanding, and principles, but only so far as they refer to objects that can be given to the senses, and thus confirmed by experience” (see [269]).

Of course, ontology should have been the science of the most general properties of real things, but the problem is precisely to know what they are. In the XIXth century, some logicians like Brentano (1781–1848) or Meinong (1853–1920) tried to answer this question and to determine on what conditions an entity must be viewed as an existing one. One of the main questions was: can we say that objects (or quasi-objects) defined by contradictory properties have some kind of existence or not? At the end of the XIXth century, the logician Twardowski considers, as Meinong and, before him, Suarez, that everything which is not nothing, but which, in some sense, is “something” (even a simple *ens possibile*, or an *ens rationum*), is an object (see [484]).

In this context, the great philosopher Edmund Husserl, in his *Logical investigations* (see [254], 28), inspired by the mathematical works of Klein, Lie, Cantor and Riemann, described the project of a new *mathesis universalis*, a science including the sum total of the formal *a priori*, whose categories should be divided into signification-categories and formal-ontological ones. This science, very organized in itself, contains different regions like formal ontology (another name for logic, the eidetic science of any object whatever), ontology of natural science, or other ontologies of the same type (see [379]).

As Husserl explains (see [255]), the categories of the analytic region of formal ontology are any object whatever:

1. A small stock of immediate or “fundamental truths” which function as “axioms” in the disciplines of pure logic;
2. The fundamental concepts of pure logic which occur in those axioms;
3. The concepts by means of which, in the total set of axioms, the logical essence of any object whatever becomes determined;
4. Finally, the concepts which express the unconditionally necessary, and constituent determinations of an object *as object*, or of anything whatever in so far as it can be something at all.

We must add that, for Husserl, one discipline not mentioned above, should have played a particular role. Its name is “apophantic logic”, which, “though it makes statements exclusively about significations, is nevertheless part of formal ontology in the fully comprehensive sense”, because one must set the signification-categories apart as a group by themselves and contrast them with the others as the formal objective categories in the pregnant sense. Later on, Husserl will oppose formal apophantic and formal ontology in a more strict sense, the former dealing with pure significations, the later being in fact an “object oriented” discipline.

Such a viewpoint may seem a very abstract one, and, it is true it is so. After Husserl, however, logic became more and more important in the field of language analysis, and different views on formal ontology were proposed.

Formal Ontology After Husserl Many kinds of formal ontologies have been developed after Husserl’s.

Lesniewski’s system of logic, for example, consists of three theories, which he calls *Protothetic*, *Mereology* and *Ontology*. If Protothetic is the most comprehensive logic of propositions, then Ontology is the most comprehensive logic of names, motivated by the fact that logic, for Lesniewski (see [309]), must be an *interpreted system* in relation with the world. As for Mereology, it is probably one of the most comprehensive pseudo-set theories, a particular system dealing with part and whole relations and where the classic membership relation has been replaced with the relation “to be a concrete fragment of”.

After him, different authors have built up other interpreted systems of logic. For example, Sommers defines “ontology” as the “science of categories” (see [467], 351), and Grossmann (1931–2010) goes on by saying that ontology asks and tries to answer two related questions: What are the categories of the world? And what are the laws that govern these categories? So, ontology is not a science among sciences. In ontology, the fundamental laws describe the behavior of categories (see [214], 3–5).

More precisely, for Nino Cocchiarella, formal ontology connects logical categories—especially the categories involved in predication—with ontological categories, the goal of it being the construction of a kind of leibnizian *lingua philosophica*, or *characteristica universalis*, as explicated in terms of an *ars combinatoria* and a *calculus ratiocinator* as part of a formal theory of predication. In this way, a formal ontology should serve as the framework of a *characteristica realis*, and hence “as the basis of a formal approach to science and cosmology” (see [93], 23). It should also serve as a framework for our commonsense understanding of the world. In this sense, ontology, for Cocchiarella as for Lesniewski, is not an uninterpreted calculus, as logic is, but a “method of constructing abstract formal systems subject to varying interpretations over varying domains” (see [92], 640–641).

Of course, there exist different forms of ontology to be compared. Since Brentano and Meinong, a very disputed question is: can there be things that do not exist? Or also: is being the same as existence? Different formal ontologies will answer these questions in quite different ways. The distinction between “being” and “existence”, where, by “existence” one means “physical existence” and by “being”, “physical

existence in some possible world”, leads to the philosophical opposition between *actualism* and *possibilism*. According to possibilism, there are objects that do not now exist but could exist in the physical universe, and hence, being is not the same as existence. On the opposite, actualists think that being is really the same as existence (see [94], 105–106).

In this context, the relation between logic and ontology may be defined in the following way. Though some authors (like Gödel, for example) think that logical form can be used not only to represent logical validity, truth conditions, abstract calculus, or cognitive aspects of our representation of the world, but even ontological structures, most of philosophers believe it is reasonable to think that the connections between the ontological and formal levels are much more complicated. An example, taken from Aristotle and quoted by Anscombe and Geach (see [7]), will show it more clearly than a long theory. Let us consider the following sentence: “The road that leads from Athens to Thebes is the same road that leads from Thebes to Athens, but in the former case it goes uphill, while in the latter it goes downhill”. There is obviously a relation between the object and the point of view of looking at it. One will say that there is two *states of affairs* for the same *situation of affairs*. Other examples may be given, like: “Heidi is Hans’s wife; Hans is Heidi’s husband”, or the famous one: “The glass is half full; the glass is half empty”. As Roberto Poli (see [400]) has shown, Husserl constructed states of affairs from situations of affairs on the basis of some rough principles:

1. States of affairs are categorical structures;
2. Situations of affairs are precategorical entities;
3. States of affairs are founded upon situations of affairs;
4. Different states of affairs may be founded upon the same situation.

Unfortunately, as Smith (see [462]) remarks, the principles Husserl put forth do not tell us what situations are (are they species, types of states of affairs, or some other things?), the states of affairs being only, in principle, *aspects* of situations. But a situation might be viewed as the matter (Sache) from which different states of affairs are formed; or it might be viewed also as a certain sort of part-whole complex, or affair (Sache), from which parts are extracted and put together into the so-called “states of affairs”.

None of the differences we may find between “situations of affairs” and “states of affairs” are totally indisputable. Maybe the only objective remark we can make about the two sentences “the road that leads from Athens to Thebes” and “the road that leads from Thebes to Athens” is that their interpretations are explicitly linked to a direction, so, correspond, from a mathematical viewpoint, to directed graphs, while the invariant of these two states of affair is that the towns (Athens and Thebes) are connected by a unique road (which corresponds to an a-directed graph). We may therefore distinguish between two representational spaces: a space composed of situations of affairs (a-directed graphs), and a space composed of states of affairs (directed graphs).

Many other examples could be given, which lead to the conclusion that, on analyzing a number of classic situations of categorial duality, we have seen the possibility of an underlying precategorical unification (the relationship between states

of affairs and situations of affairs). The unification therefore acts as a quotient or invariant. Then, if it were possible to generalize the procedure by unifying a multiplicity of unfolded states of affairs in their underlying state of affairs, we should have access at some quasi-objective ontology.

But in fact, duality phenomena seem to be a particular case of a more widespread phenomenon, difficult to formalize. Besides pairs of categories with a recognized mathematical meaning, like finite/infinite or discrete/continuous, one easily finds many further ones that have figured prominently in the history of philosophy. For example, the opposition between matter and form, potential and act, quality and quantity, one and many, identity and difference, individual and universal, part and whole, and so on. Then, a further investigation shows that the dualities that hold within these pairs of categories are not elementary: the form of a certain substance is the substance of another form; the act of a certain potential is the potential of a new act, etc. So, the categorical duality that holds in ontology is not necessarily a simple one. And therefore, building a formalized ontology means to suppose ontological assumptions that impose, indeed, some constraints on the world of referents.

With the above views, the relation between ontology and classification has begun to reveal its importance: the main categories of a classification—for instance, a classification of knowledge or a library classification—could highly depend on semiotic oppositions or categorical dualities like those listed above⁴ and connected to particular interpretations of some situations of affairs. Ranganathan's classification, for example, has been viewed as directly linked to some of the main categorical oppositions of Indian thought. It seems that it is difficult, and probably too ambitious indeed, to pretend, by these means, achieving some universals.

Ontology should be, in principle, the theory of “what there is”. And though we can make distinctions, in a heideggerian style, between ontics (i.e. a purely descriptive and analytical discipline), and ontology (a more speculative and formal one) (see [392]), ontology is generally defined as a “theory of objects”. As Roberto Poli says, it may be even “a theory of every type of objects concrete and abstract, existent and non-existent, real and ideal, independent and dependent” (see [399]). So, the problem is now to explain what are the objects of this theory and how they are organized.

Since the origin of philosophy in Ancient Greece, there has been different ways of solving such problems. In general, objects have been defined according to their

⁴The existence, in the language, of such oppositions, and particularly the difference between contradiction and contrariety, has suggested to Sommers (see [468], VII–VIII), a staunch proponent of a traditionalist view of logic, to develop a new “Calculus of Terms”, significantly different—in his eyes—from the predicate logic, but closer to some of the Leibniz's proposals. In particular, in [468], Chap. 13, the author indicates how traditional logic's way with contrariety leads to the conception of categories that is at the basis of Ryle's seminal work in the forties (see [438]) and his own more formal treatment of categories in the early sixties (see [467]). More precisely, Sommers recognizes the need for a notion of contrariety that would allow for saying, for example, that “Saturday is neither fed nor unfed” (which renders both “Saturday is fed” and “Saturday is unfed” *category mistakes* in the sense of [438]). This kind of problems prompts him to re-examine traditional Aristotelian logic and its characteristic distinction between contrary terms or predicates and contradictory propositions.

properties and one used to think that a domain is well organized when objects can fall in some type or category. But there are in fact many views on properties, some of them being explored by Alex Oliver in a famous paper (see [369]). As the author shows, properties may appear as sets of particulars, as universals, as sets of tropes, etc. and, after years of investigations, many questions are still open. Must we accept, for instance, Lewis' natural properties or Armstrong's "states of affairs"? The answer depends widely on a philosophical choice. In the same way, there is no clear understanding of what precisely an ontological category is, and recent developments on the subject rather defend the idea that there is no unique system of such categories, the ontological category an object belongs to being not an essential property of that object. As a solution to the old problem of universals, must we prefer Carnap's scientific realism (see [77]) to Goodman's amendment of it (see [199], 135–155) or must we accept only some kind of "resemblance nominalism" (see [425])? In this context, systems of ontological categories seem to be particular structures imposed on the world, rather than reflections of a deep metaphysical reality already present (see [498], and the analysis of [73]).

As there is no consensus, between philosophers, to define "what there is", finally, as Kit Fine said, "an ontology consists of all those items which are, in an appropriate sense, accepted" (see [169]). And, of course, there are different views as to what it is for an item to be accepted into an ontology. But whatever the (philosophical) choices might be, an item is accepted into an ontology because, in the last analysis, *it should be there*. It is the ontology which accepts the item, not the person who endorses the ontology. So the rules we must construct for that are very important.

Moreover, to be an effective construction, a general formal ontology must apply, and so, it may be useful to divide it into different domains: which means that one should have to build in fact, for the concrete world, *regional ontologies*.

2.8 Regional Ontologies

Building a Basic Formal Ontology (BFO) is now a big international project, of which Barry Smith, the Coordinating Editor of the OBO⁵ Foundry and a member of several other working groups in ontological biomedical investigations, is the leader, and which spreads in different domains of knowledge (biology, biomedical science, geography and geospatial studies, social and cognitive sciences). In all those domains, Barry Smith and his group try to construct formal ontologies of objects, functions, categories, processes, acts, etc., which tend to structure these domains in a satisfying way. We shall start by explaining the necessity and relevance of such a view, then we shall study the philosophical assumptions of Barry Smith's approach, and finally discuss some aspects of his very interesting attempt.

⁵Open Biological and Biomedical Ontologies.

2.8.1 The Need for a “Metaclassification” Approach

In many domains of science today, there are a number of reasons why some epistemological viewpoint about classifications must be performed:

1. No classical ontology explicitly explains the complex use of names in scientific abstract sentences. See, for example, in biology, the following one: “DNA-binding requirement of the yeast protein *Rap1 p* as selected in silico from ribosomal protein gene promoter sequences” (this example has been given by Barry Smith in some of his courses in Buffalo).
2. Use of different classificatory aspects rests on tacit specialist knowledge, and there is no (automated) method to derive such a secret and mysterious science.
3. “Is-a” and “part-of” relations have often been employed by scientists in inconsistent ways.
4. Some relationships can be verified automatically, but where entities have been already classified in a way which hinders metaclassification, one needs to modify the relations (manually).

As Smith and Kumar say in the conclusion of a paper read at the Symposium on Bioinformatics and Bioengineering of Taiwan (see [460]), there is a chance that such problems exist with most, if not all, biomedical ontologies and classifications, and therefore, a fundamental review is needed in order to address more pertinent issues. The hope of the authors is that, after a period where the application of pragmatic principles that have led biologists and biomedical researchers to the creation of relatively simple models and have facilitated the storage of large amounts of biomedical data, then a next era must open up to new ideas. In particular, the adherence to formal principles of representation, which, for the authors, is supposed to make such data better manageable, could contribute to more robust bioinformatic science.

Of course we share Smith’s opinion that a deep and strong metalevel theory is necessary before using of clustering algorithms or running computer softwares. But the question is: what must we take as “metaclassification” principles? After having investigated the method and results of Smith, we try to explain our point of view.

2.8.2 Barry Smith’s Project

The basic idea of Barry Smith is that, after Aristotelian realism and Kantian constructivism, the third metaphysics which has dominated the XXth century is the theory based on advances of predicate logic. Its syntax is inspired by the symbolism of function and argument. For example, there are atomic sentences like $F(x)$, $R(x, y) \dots$, and molecular sentences a little more complex like $(F(a) \wedge G(b))$ or $F(a) \supset \exists x, R(a, x)$, or again $\forall x (P(x) \supset \exists y, L(x, y)) \dots$. As a result of the works of Frege, Russell and Wittgenstein, this syntax came to be awarded a special role in the practice of philosophy. As reality may be decomposed into objects

and properties (or attributes), so that we can get matrices indicating what object possesses (or not) what property, and we spontaneously apply formal logic to the description of the world. In his lectures, Smith often calls “Fantology” (see [461]) the doctrine, usually tacit, according to which “ $F(a)$ ” (or, of course, $R(a, b)$, etc.), is the key to the ontological structure of reality. But there are many differences between formal logic (dealing with interconnection of truths) and formal ontology (dealing with interconnection of things). Logical structure is not ontological structure. For example, “entails” is a logical relation, while “part-whole” is an ontological relation. So, a first mistake of “fantology” is that all form is logical form: as he says, “the fantologist sees reality as being made up of atoms plus abstract (1- and n -place) ‘properties’ or ‘attributes’.” But it is not the way scientists use names. A simple glance at the regional ontology of a particular domain (say, for instance, proteins in biology) shows that there is an ontological complexity. Generally speaking, for Barry Smith, fantology reductionism cannot do justice to the multiple levels of “granularity” of reality. The logical atomic, timeless structures cannot be connected up with the real entities (which are, in different ways, associated) until, in some future state of “total science”, an hypothetical perfected physics of ultimate atoms reveals finally the true ontology of the world. The fact that the same form $F(a)$ is used to express more or less any kind of content leads to other mistakes. For example, one tacitly assumes that to every predicate a property corresponds (but this only means that such a doctrine of properties is ontologically empty). Similarly, the “fantologist” believes that there is only one state of affair. But the absence of distinction between, for example, predication in the category of substance and predication in the category of accident (a is an elementary particle, a is negatively charged) brings a confusion of universals and properties. Moreover, Fantology based on boolean structures, cannot represent changes of time, so forgets all processes, and is conducted to present a poor treatment of relations, environment, family resemblances, etc. Finally, it has so much limitation that it rapidly leads into the temptations of possible worlds metaphysics.

Mathematics is also included in this list of grievances. For Barry Smith, in several points, mathematics does not go further than logic, particularly in the treatment of sets and wholes. A set in the mathematical sense was initially conceived by Cantor as “a collection into a whole of definite and separate objects of our intuition or thought”. But set theory has since departed in several ways from this conception, so that, now, groups and wholes, as concrete denizens of reality, are very different from sets. For example, there is nothing in the realm of concrete groups or wholes analogous to the empty set, no monsters like $\{\emptyset, \{\emptyset, \{\emptyset, \{\emptyset, \{\emptyset, \{\emptyset\}\}\}\}\}, \{\emptyset\}$, no “pure sets”, no stable entities compounded out of members or parts arbitrarily. Real groups and wholes, like all entities in physical reality, are subject to change and, above all, may gain or lose members while preserving their identity. Finally, Smith makes a difference between two kinds of groups: “bona fide groups”, which exist in reality independently of human cognition (for example, the planets of the solar system, the group of rabbits on the Island of Gozo, the species cat, the genus mammal) and “fiat groups” which exist only as the product of human fiat, and thus of some cognitive process (for example, the European Commission, rabbits in Trentino, the Republican voters in Dade County, the Finnish diaspora). For Smith, “there are no fiat

groups in the extra-human world” and it is thus “a sad reflection on the exact philosophy of our day that its two principal formal tools—set theory and mereology—apply, at best, to counterparts of fiat groups. A formal theory adequate to the realm of bona fide groups does not yet exist” (see [461], 130).

2.8.3 *Mereotopology and Granular Partitions*

So, for Barry Smith, an alternative to “fantology”, might be set out for a better applying of mathematical concepts to the real complexity of the world, at different levels and in different domains: that is what the author tried to do. In more that 450 publications, 15 books and the management of impressive projects, the author has begun to build formal regional ontologies. He has done all that with the help of eminent scientists in several domains like medicine, biology, genetics, geography or ecology, and our intention is not to criticize a so massive work.

However, from a mathematical viewpoint, the reform of the father founders’ logic and of classical set theory that Smith intended to perform, has been pursued essentially in two ways: first by constructing a concrete mereotopological theory of parts and wholes, and second, by trying to adapt mathematical concepts (for example, the notion of partition) to the needs of an ontology of real objects. Let us say a few words about these two attempts, on which, maybe, some reservations may be expressed.

Mereotopology The construction of a mereological viewpoint in ontology began with a paper devoted to Chisholm (see [456]), where Barry Smith deals with the treatment of boundaries, a category of entities usually neglected in ontology. Spatial bodies are concrete continua with boundaries and a naive would think that they may receive a mathematical treatment. But for Smith, the set theoretic approach to the continuum is inappropriate to understand those structures. Some bergsonian style arguments have been put forward: concrete continua are qualitative and heterogeneous structures with multidimensional boundaries where nothing like the Cantor’s problem may arise, while the set theoretic continuum is supposed to be made of homogeneous isolated basic elements or blocks, problematically assembled together to form a whole of higher dimension. Following Chisholm, who has adopted mereology, Smith considers boundaries of bodies to be actual parts of the bodies which they bound. In this way, boundaries can be external, internal, asymmetric and need to be classified. Following Brentano, he admits that points may have parts (called “plerotic parts”), and proposes a curious theory of coincidence where, as he said, bodies do not coincide: not even with themselves, nor do they with the spatial region they occupy (which, of course, contradicts the identity principle). In this theory, boundaries are finally defined as plerotic self-coincident proper parts of the bodies. As the theory, though a pseudo-mathematical style of presentation, has not been axiomatically seriously introduced, it is difficult to test its consistency or its soundness. Smith recognizes himself that some spatial “monsters” could call some axioms into question, and he hardly avoids some undesirable consequences of them.

Granular Partitions Nevertheless, the ideal of building a mereotopology is pursued in another paper where the question is to solve problems which arise when mereotopological methods “are extended to deal with those varieties of spatial and non-spatial reasoning which involve a factor of granularity” (see [457]). Let us make this question precise.

Mereology, as region-based, yields a more realistic representation of qualitative space of common sense. However, while factors of vagueness, imprecision and uncertainty could have been added to standard mereotopological approaches, mereotopology resists a similar realistic extension to a theory of granularity. In mereology, when an object falls within the range over which one quantifies, then also do all the object’s parts. In Set theory, the axiom schema of separation blocks the automatic recognition of an object’s parts, but, according to Smith, only at a too high sacrifice in realism.

To organize the relations between an object and its location in a more realistic manner, Smith introduces the concept of “granular partition”, which is very different from the mathematical concept of partition (classically defined here in Chap. 3 in terms of equivalence classes), and which should have been, in fact, called by another name, because it is not a partition at all. Roughly speaking, Smith’s partitions, as *systems of cells*, are grid-like representations at different levels, that project outwards towards those objects in reality which are located in their cells. Let us consider, first, the formal definition of those systems (as far as we can reconstitute it from [457] and [53]):

Definition 2.8.1 (Smith’s granular partition) A granular partition A is a finite system of cells $(z_1, z_2, \dots, z_i, \dots, z_n)$ on which is defined a “sub-cell relation” $\subseteq_A$, which is a reflexive, antisymmetric, and transitive relation.

Theorem 2.8.2 A granular partition A has the following properties:

1. The transitivity of $\subseteq$ generates a nestedness of cells inside the partition in the form of chains of cells $Z = \langle \dots \supset_A z_i \supset_A z_{i-1} \dots \rangle$ having a maximal element $z_{\max}$ and a minimal element $z_{\min}$;
2. If two cells in the partition overlap, one is a sub-cell of the other: $\exists z: (z \subseteq_A z_1 \text{ and } z \subseteq_A z_2) \implies z_1 \subseteq_A z_2 \text{ or } z_2 \supset_A z_1$;
3. The partition theoretic union $z \cup_A z'$ in the partition A is a $\subseteq_A$ -minimal cell (called the root) satisfying the condition that it contains z and z' ;
4. The partition theoretic intersection of z and z' in the partition A (or the overlap of z and z' in A) is any $\subseteq_A$ -maximal cell, included as a sub-cell within them both, and defined by: $z \circ_A z' := \exists z (z \subseteq_A z_1 \wedge z \subseteq_A z_2)$.
5. Each cell in a granular partition A is connected to the root via a finite chain of immediate successive cells: $\forall z_1, z_2$ in A , if $z_1 \subseteq_A z_2$, there is no z_3 such that $z_1 \subsetneq_A z_3 \subsetneq_A z_2$.

Remark 2.8.3 If A were a classic set, $\subseteq$ would define on A a large order relation, so A will be just a poset.

Remark 2.8.4 Which Smith says in condition 2 is the exact opposite of what is generally understood under the word “overlapping” in classical clustering analysis all over the world: for Smith, overlapping reduced to classic inclusion.

Remark 2.8.5 Partition theoretic union and intersection (condition 3 and 4) are commutative but not associative and, in general, not unique. So they determine a structure which is close to what we called, in Chap. 4, “non-associative product”, though no algebra is defined on it.

Remark 2.8.6 Condition 5 involves in fact two assertions: first, the fact that A is connected, and second, the postulate that a chain Z of successive cells is necessarily discrete and nowhere dense.

In the following, Barry Smith explains finally that his “granularly partition” means in fact nothing else than a rooted tree of finite depth. The connection between partitions and trees will now be obvious: it is a simple matter to show that every finite partition can be represented as a rooted tree of finite depth. The method for constructing a tree from a finite partition consists of the following trivial steps:

1. Creating a graph by mapping the cells z_i onto nodes v_i within the graph;
2. Introducing a directed edge from vertex v_i to v_j iff the cell z_i has z_j as an immediate sub-cell.

This is always possible, as it follows from the fact that the sub-cell relation is well defined and that chains of immediate cells are always finite (see Definition 2.8.1). We can also easily show that the resulting graph is a rooted tree (which can be deduced from Theorem 2.8.2(3)), that the graph structure is connected (Theorem 2.8.2(2)), and acyclical (2); and that there is a unique path between any two vertices (from Theorem 2.8.2(5)). The complementary reconstruction of a “granular partition” from its tree representation would be no less trivial.

In another part of his papers, Barry Smith then tries to face up to the problem of the relations between our grid-like models and the changing aspects of the world’s objects, which are not always located in the same place at different times. So, the formula $L_A(o, z)$ will abbreviate: object o is located in the cell z of the granular partition A . For Smith, such a location L presupposes a “projection” P . As he says, “an object is never located in a cell unless the object has already been picked out as the target of the projection relation associated with the relevant partition” (see [53]). So Smith defines the condition of successful projections (or, vice versa, of successful locations) which yields “transparent partitions”. So we have the following properties for projections:

1. $L(o, z) \longrightarrow P(z, o)$;
2. $P(z, o) \longrightarrow L(o, z)$;
3. $P_A(z, o) \longrightarrow L_A(o, z)$ (transparency of A);
4. $P(z, o_1) \text{ and } P(z, o_2) \longrightarrow o_1 = o_2$;
5. $L(o, z_1) \text{ and } L(o, z_2) \longrightarrow z_1 = z_2$;

6. $\forall o_1, o_2, L(o_1, z_1) \text{ and } L(o_2, z_2) \text{ and } z_1 \subseteq z_2 \longrightarrow o_1 \leq o_2$ (representation of mereological structure (RS));
7. $\forall z_1, z_2, Z(z_1, A) \text{ and } Z(z_2, A) \longrightarrow RS(z_1, z_2)$ (mereologically structure-preserving partition);
8. If $r(A)$ is the root cell of the partition A , then $\exists x: P(x, r(A)) = D(A)$ (a partition A has a non-empty domain $D(A)$).

If conditions (1)–(8) are satisfied, then the granular partition is said *to represent its domain correctly*. So, Smith goes on by studying structural properties of correct representations (mereological monotony, completeness, exhaustiveness) and the axioms that must be satisfied in this case (such as the comprehension axioms). But the main reasons for introducing such concepts are the other situations: case of redundancy, incompleteness as relative fullness or cumulativeness.

Redundancy will happen, for example, in the following case: Assume there exists a partition with a cell labeled “vertebrates”, which occurs as a sub-cell of the cell labeled “chordates” in the standard biological classification of the animal kingdom. Actually, we know that almost all chordates are in fact vertebrates. But suppose (for the sake of argument) that biologists were to discover that all chordates must be vertebrates. Then a redundancy will appear in the classification. So, such a discovery would imply that, in order to avoid this structural redundancy, one would need to collapse into one cell the two cells (of chordates and vertebrates) which at present occupy distinct levels within their zoological partitions.

In the case of incompleteness, some examples of non-full and non-cumulative partitions as those which are represented on Fig. 2.2, will be more enlightening than words.

Figure 2.2(a) is a partition full and non-cumulative; for example, the belief of some child who thinks that cats and dogs are the only animals there are.

Figure 2.2(b) is a partition non-full and cumulative. It is the way a child sees the world who does not understand the concepts of the Northern and Southern hemisphere and who thinks that there are places on Earth that are neither in the Northern, nor in the Southern Hemisphere, nor overlapping both, but are rather in some secret and wonderful land that has not yet been discovered.

In Fig. 2.2(c), we have a non-full and non-cumulative partition. For example, the partition you build when you think of last night’s party as consisting of a root with three sub-cells: you, John, and Mary, but when you know you are missing someone.

In conclusion, variant forms of Smith granular partitions seem to have been advanced, in particular in the literature on Spatial Information Science. No doubt that they may be helpful in geographical sciences, especially when vagueness is taking an account: for example when one deals with the definition of some ontological object whose proper name, say, for instance “Mount Everest”, refers to some mereological whole (a certain giant formation of rock, as it happens) with more or less indetermination in location (see [51]). No doubt that they throw new light in the field of formal (or formalized) ontologies of some concrete domains.

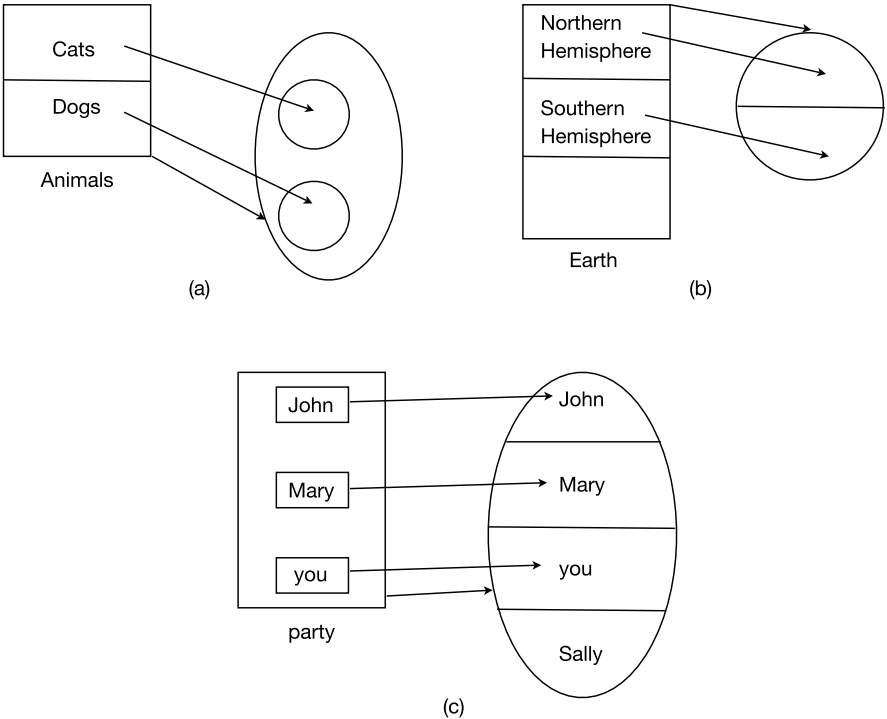


Fig. 2.2 Examples of non-full and non-cumulative partitions

2.8.4 Discussion

However, we must say that Smith’s concepts, in general, are probably better adapted to solve formal ontological problems than classification problems.

From the viewpoint of a classification theory, we are not absolutely convinced of the usefulness of Smith’s approach. We are even inclined to think there is a chance this language introduces a certain confusion in subjects that have been well described in mathematical literature, with the help of the usual mathematical concepts, which are not so bad: in this way, the famous “fantology” is, for the essential, a fantasy of the author: mathematics does not reduce to predicate calculus, nor to boolean structures and a lot of possibilities for treating changes and time exist in mathematical physics, from phase spaces to Lie groups.

Of course, the obvious limitations of set theory have often generated transgressive projects. Lesniewski’s mereology is one of the best examples of such theories. But Smith’s attempt of associating topological concepts to a mereological system is quite unusual, and a bit surprising for a logician. In 1991, Parrochia (see [377]), following some remarks of Laborde (see [289]), has shown that mereology, in itself, could not yield a topology and was in fact connected with pre-topological (or pseudo-topological) invariants like the notion of “pseudo-boundary” which allows

us to think of overlaps defined as superpositions of sheets (for instance two cultures on both sides of the frontier between two countries) and not as intersections of sets. Representing this kind of phenomena would be, among others, a good reason to use this theory. But mereology is not a good system for representing topological situations.

We can doubt, in fact, that the introduction of mereology can be very relevant in general for a classification theory. On the contrary, using mereology raises a lot of problems. As mereological classes are collective classes and not distributive classes, the definition of a domain like “dogs” in the animal kingdom, for example, implies that, “besides the species dog, also your dog Fido, and also Fido’s DNA-molecules, proteins, and atoms are parts of the domain of this partition” (see [53]). But as the latter are, of course, not recognized by the partition itself, mereology requires supplementation by a theory like the theory of granular partitions. But taking granularity into account in a theory (which, usually, has nothing to do with it) requires a definition of a restricted parthood relation, which is an analogue of partition-theoretic inclusion, but far more complicated. Is this sufficiently motivated to be done?

Another problem raised with Smith’s approach is the following one: it is very doubtful that the concept of “granular partition”, in the case when the projection is correct, says far more than what is already known under the best-chosen name of “classification” in a classic sense (a viewpoint that has been already described for a long time by all mathematicians around the world within the frame of classical order theory).

We can also observe that the taxonomy of these “transparent” partitions, which results in the classic case, is not founded on structural properties of partitions themselves but on some poor quantitative concepts of “degree”. Smith classifies granular partitions along three essentially orthogonal axes: (a) degrees of correspondence to objects; (b) degree of structural fit; (c) degree of completeness (see [52]). But this left open the basic problem of a classification of classifications based on their real structures and transformations.

Of course, we can easily agree that mereology shows some advantages: for instance, it does not make use of empty sets (but this raises also some other problems because Smith has to invent a complex definition of an empty space for using his “empty cells”). Another important point of Barry Smith’s approach is that we can make outwards projections of the grid-like so-called “granular partitions” in different ways. But, if we except these few innovations, it seems to us that the mathematical model proposed by Smith brings nothing new and generates a big complexity which is not always useful.

Finally, we contest that we must relax in such a way the relations between our mathematical models and the real world, to which they apply, the most of the time, without unreasonable idealizations.

In fact, for a best understanding of the viewpoint of this book, we ask the reader to consider the schema of Fig. 2.3.

A simple glance at the history of classifications proves that some advances happen in classification theory when phenomenological classifications are replaced with what we called “noumenal classifications”, i.e. when clustering methods with more

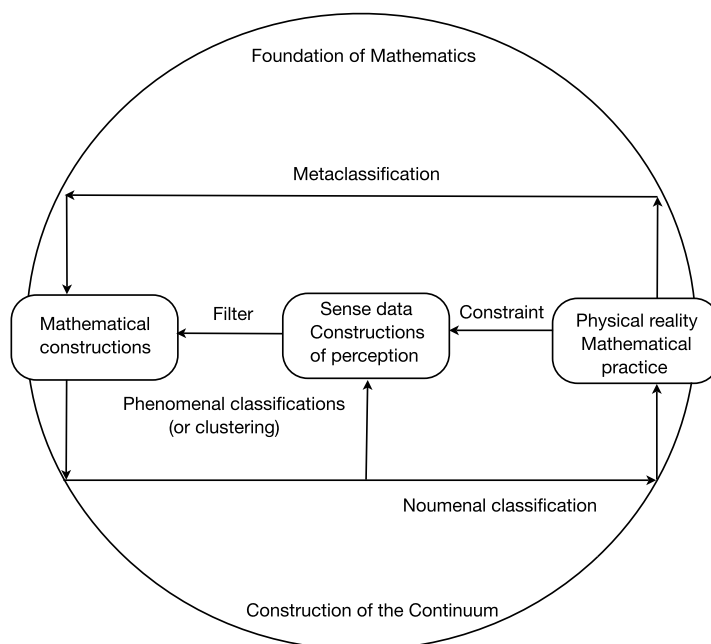


Fig. 2.3 An overview of the classification domain

or less conventional dissimilarity measures between pairs of objects essentially based, in last analysis, on sense data, are replaced with good mathematics describing the intrinsic structure (or morphology) of these objects. But, at their turn, finite or infinite structures of this kind are deeply dependent on a metaclassification view which points out the construction of the continuum as the infinite set of all classifications, a metastructure at least standing as a “regulating horizon”—as Kant would have said—but whose properties are progressively investigated by mathematicians. A general theory of classifications would be the construction of the continuum and, equivalently, a foundation of mathematics.

2.9 Ontologies, Formal Concept Analysis and Data Mining

All the domains we have explored in this chapter belong in fact to what is called Data Mining. Ontologies are connected to Formal Conceptual Analysis, because, as we know, ontologies are explicit specification(s) of a conceptualization (see [216]).

In this sense, they usually consist of a set of concepts (not to be confused, however, with “formal concepts” from FCA, even though there are some links between them), a hierarchical “is-a” relation, some other (non-hierarchical) relations between the concepts, and eventually axioms describing constraints on the relations and concepts.

These problems arise when we have to constitute the space of data on which we have to work. Today, very often, this is done automatically. For example, one of the main tasks, in learning ontologies from data, is the construction of the “is-a” hierarchy.

Suppose that the concepts are already learned, e.g., by applying linguistic and statistical methods (see [326]) and stored in the set M . The set G contains “tuples” of a relational database, or documents annotated with the concepts. The relation indicates if a tuple includes a concept, or if a document is annotated with a concept.

For example, TITANIC, one of the methods using weight functions, can assign to a set X of ontology concepts the number of documents/tuples related to all concepts in X . The resulting iceberg concept lattice (see section above) provides an “is-a” hierarchy on the set of the ontology concepts. Additionally, it can also suggest new concepts which may simplify the structure of the concept hierarchy.

But, though a good classification can also be used to discover new objects and make science progress, this kind of problems is quite different from the main problems of a general theory of classifications. We have indeed to progress in the direction of an algebra of classic, and, if possible, generalized classifications.

Towards a General Theory of Classifications

Parrochia, D.; Neuville, P.

2013, XXIII, 304 p. 72 illus., 1 illus. in color., Softcover

ISBN: 978-3-0348-0608-4

A product of Birkhäuser Basel