“Semiotic Aspects of Generalized Bases of Data”, in E. Kawaguchi, H. Kangassalo, H. Jaakkola, I.A. Hamid, eds., *Information Modelling and Knowledge Bases XI*, IOS Press, Amsterdam, 2000, pp. 1-11

Semiotic Aspects of Generalized Bases of Data

Roberto Poli and Guerino Mazzola
Department of Sociology and Social Research – University of Trento, Italy
Instituk für Informatik / Multimedia Lab – University of Zürich, Switzerland
roberto.poli@soc.unitn.it; guerino@mazzola.ch

Abstract. We have systems rich in information but poor in knowledge. The paper shall sketch a framework in which information and knowledge could be properly integrated. Our initial hypothesis is that passing from information to knowledge involves the elaboration of a well structured theory of semiotic units. The latter is codified by a three-fold structure: an expression, a content and an ontological component. Expressions and contents together give rise to the symbolic component of the semiotic unit. The three components are all further subdividable into types and kinds. Their intertwined relationships become easier to elaborate if we find a uniform way to represent the various components. This requires formal tools as sophisticated as those provided by category theory. It follows that other operators beyond the usual logical operators should be used, namely the 'geometric operators' of category theory.

1 Introduction

Unlike the societies of the past, modern ones are no longer afflicted by a lack of information. If anything they suffer from its excess, from having to cope with too much unused and unusable information. It becomes increasingly difficult to find the information that one needs, when one needs it, to the extent that one needs it and in the appropriate form. Although the information may be stored somewhere, all too often one does not know where; and even when one is aware of how to find the information, it is often accompanied by further information irrelevant to one’s purposes. And when information is available, it is often forthcoming in the wrong form, or else its meaning is not explicitly apparent.

However broad the range of information already gathered may be, a great deal more has still to be assembled and codified. And this inevitably complicates still further the problem of the functional, flexible, efficient and semantically transparent codification of information.

The problem is that we have systems rich in information but poor in knowledge. We are in need of tools able to connect information to the events and objects of the world; instruments able to process, translate and deliver information in forms suitable for the problems to be addressed and solved. These are problems of knowledge, not merely of data theory.

In the following we shall sketch a framework in which information and knowledge could be properly integrated. Let us start with some preliminary clarifications.

Following on [3], we shall assume that a *symbolic unit* is given by an *expression unit* plus a *content unit*. Let us write a symbolic unit in the following way

$$\frac{\text{table}}{\text{TABLE}}$$

where ‘table’ is the expression unit and ‘TABLE’ the content unit. Needless to say, we can have both

$$\frac{\text{Tisch}}{\text{TABLE}} \qquad \frac{\text{tavolo}}{\text{TABLE}}$$

that is, different expression units for the same content unit (in this case, pertaining to different natural languages) and

$$\frac{\text{bank}}{\text{BANK1}} \qquad \frac{\text{bank}}{\text{BANK2}}$$

where 1 and 2 can be taken to be contextual indexes (concerning the above exemplifications, 1 = river, 2 = financial institution).

Expressions can be written as linguistic marks (as in the examples above), logical and mathematical formulas, acoustic marks, musical, pictorial ones, etc. Linguistic marks obviously concern not only written signs but also acoustic ones, as in

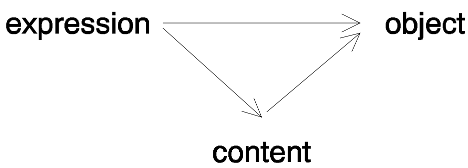
$$\frac{\text{'teibl}}{\text{TABLE}}$$

(here represented in the form of a written sign. Note that the expression units for the four classical types (string, floating, integer, boolean) are all given as strings. Some sections of the other possible types may be of the same nature (i.e., the phonetic alphabet), but others may be very different.)

For symbolic units (that is, unities of both expressions and contents) it can be distinguished between symbolic units that have *referential capacity* and symbolic units that do not have referential capacity. As far as linguistic expressions are concerned, units with referential capacity are categorematic expressions (mainly open class terms), whereas units without referential capacity are syncategorematic expressions (closed class terms). Having referential capacity does not mean actually being able to refer. To know which expressions refer and which do not, we have to look at the world and the way in which it is made. Symbolic units with referential capacity always purport to refer to some kind of objectuality. They may nevertheless fail to ‘get at’ their referents, without any loss in their referential capacity.

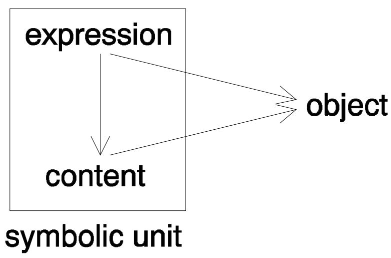
For symbolic units with referential capacity, a third component must be considered: the object referred to by the symbolic unit. We thus obtain the usual semiotic triangle:

Fig. 1



The analysis presented here suggests a slightly different picture, namely

Fig. 2



The referential capacity of symbolic units can be properly split between (1) the true referential relation between expression and object, and (2) the denotation relation between content and object. The relation between expression and content will be termed 'meaning relation'. We may therefore say that expressions *refer* to objects, that contents *denote* objects, and that expressions *mean* contents.

The *inverse* relation(s) between expression and object is/are called relation(s) of *nomination*, the inverse relation(s) between content and object is/are relation(s) of *description*, and the inverse relation(s) between expression and content is/are the relation(s) of *codification*.

In this paper – following [10, 15] – we shall propose a compositional theory of semiotic triangles following on the generalization of the classical Saussurean viewpoint suggested by contemporary cognitive semantics [6, 7].

2 Expression units

A well established prerequisite for the automatic treatment of data is that they should be systematically classified into a pre-established series of types. The most usual types are obviously those of string, floating, integer and boolean. Of course, this series is dramatically inadequate to the needs of actual bases of data. The Internet provides the most striking evidence that a new series of expression units is required. Consider the enormous amount of figural (pictorial) and acoustic signs available, and the need to organise them rationally (or at least not too inadequately).

This means that the usual series of types (string, floating, integer and boolean) should be enlarged to include at least, say, 'pixels' and 'voxels' (for digital 2D- and 3D-figural units), 'pcm-quanta' (for digital acoustic units; note: pcm = pulse code modulation), or 'phonems' (for phonological units), etc.

The point to be stressed is that expression units have to be typified. This means that the previously given structure requires some sort of integration, for instance as follows:

$$\frac{\text{table: string}}{\text{TABLE}}$$

Expressions can be natural or conventional, and conventional expressions can be conventional at different levels of conventionality (for instance, as a structured mark – word or sentence – of some natural language, or of some artificial language: mathematics, for one). In many cases expressions are more or less perfectly translatable to one another.

We have already seen that the same content can be given in a wide range of expressions, both of the same and of different types. TABLE can be given either as 'table,' 'Tisch,' 'tavola,' ecc. (all of the type 'string'), but it may also be given in the form of a sound (as in: 'teibl) or in the form of pictures, drawings, etc. (of the supposedly new type 'pixel'). Not all the types are available, however. TABLE cannot be given in the form of 'floating' (on the other hand, a TABLE'S WEIGHT can possibly be given in the form of a floating. We shall consider this aspect of the problem in § 5 below).

Types filter expressions. They specify the form in which the expressions are given. To complete the picture we should specify the *atoms* of the pertinent expression. That is to say, the alphabet for string, numerals for integer and floating, truth values for boolean. Their filters may be called sub-types and can be exemplified by latinAlphabet, as distinct from, say, polishAlphabet, or by arabicNumerals as distinct from romanNumerals, or by the truth values 'true' and 'false' as distinct from '1' and '0'.

The structure of types and sub-types also sheds light on the problem of which layer of reality is affected by a specific type. For example, strings of letters are a mental abstraction from more acoustically oriented units of the phonemic alphabet. In the field of music, the purely acoustic unit 'pcm-quantum' does not relate to musical abstraction. To grasp this layer of reality, we would rather use note units as described by classical European tradition. However, like with English-oriented phonemes (as described by the Jakobson-Halle-Chomsky system), there is no universal alphabet for mental units of music either. The problems with pixels are even more striking.

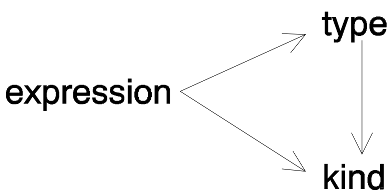
Observe that the ‘atomic’ character of type units is always relative, be it in their character as indecomposable quanta, be it in their usage in cultural contexts (See [11] and [15], ch. 5 for further details).

From what we have seen, it follows that enlarging the series of types is a major task. Yet types alone are not enough. Besides types, we may resort to what can be conveniently called *kinds*. Whereas types are intended to specify the *formal* characterization of expressions, kinds are intended to provide their *material* characterization. In the same way as types, kinds can be divided into sub-kinds, comprising codes (and sub-codes), etc. By way of example, suitable kinds for the string ‘table’ are the following (where CN is short for Count Name)

naturalLanguage : English : CN.

Therefore, the overall picture becomes:

Fig. 3



where both types and kinds are properly sub-filtered.

3 Content units

Content units can be also decomposed. The major problem here is that we do not have a well established theory concerning their decomposition into sub-complexes. There are instead compelling reasons for arguing the unsuitability of any decompositional program. However, a way out of the impasse is provided by suitable geometrical structures (we shall add some words of explanation below, leaving extensive reasons in favour of this claim for another occasion).

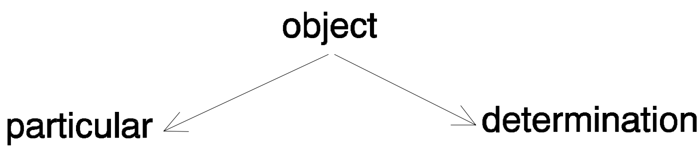
The distinctions between types and kinds for expressions and contents take slightly different aspects, because contents present only one top kind, namely *labelled descriptions*. Otherwise stated, there is only one maximally general kind of content. From a material point of view, all the contents are labelled descriptions. The canonical example of a labelled description is a dictionary entry. Needless to say, labelled descriptions can be variously sub-kinded.

On the other hand, contents units – like expressions – also present a number of types. Sets, trees, fields and the whole range of mathematical structures are types of contents.

4 Objectual units

Objects too can be typed and kinded, like expressions and contents. We may assume that the following are the major distinctions within objects

Fig. 4



The subsequent articulation in sub-types and kinds applies to determinations and particulars.

Particulars have a very complex structure comprising (1) level, (2) composition and (3) nature (see [14, 15] for details). 'Level' relates to the theory of levels of reality (material, psychological, social, and ideal, each subdivided into appropriate sub-levels). 'Composition' concerns part-whole relations. Miller's distinction between external and internal hierarchies well exemplifies our level/composition opposition [12].

The division between types and kinds articulates the nature of particulars. Kinds of particulars are at least those of stuff, individual, process, group and mixture. The first four are well recognized, which exempts us from providing support for them. As to the latter, mixtures are characterized by the intervention of a process which fuses the elements of a multiplicity together. A mixture is obtained when several ingredients cease to exist as distinct ingredients and form a (new) homogeneous whole.

Kinds can be divided into natural, dependent and hybrid. Natural kinds, in turn, are usually subdivided into kinds of individuals (tiger, narcissus) and kinds of stuffs (carbon, gold, water). Nothing prevents further articulation which comprises the other types as well. Dependent kinds can be functional (table, banknote) or contextual (cliff, bush). Hybrid kinds, lastly, comprise hybrids of natural and dependent kinds (vegetable) and hybrids of functional and contextual kinds (surfing beach, oasis, gravel pit) [19].

Determinations can be distinguished into intensive and extensive. Extensive determinations

- are extended (they always exist between two points);
- for every determinable there is only one determined instance;
- the instantiated attribute of the whole contains smaller parts (diverse as instances) of its determinable (volume in volumes, shape in shapes);
- are never lower-dimensional;
- can be summed (aggregated);
- have a prothetic order (based on 'more than').

Complementarily, intensive determinations:

- are not necessarily extended (they may be punctiform in nature);
- for every determinable, a whole may have many determined instances;
- do not contain smaller parts (diverse as instances) of the determinable;
- may be lower-dimensional attributes of other attributes;
- cannot be summed (aggregated) (the colour of a whole is not the sum of the colours of its parts, if anything it is the pattern of the colours of its parts);
- have a metathetic or positional order (based on 'different from').

The characteristics of extensive and intensive determinations are taken from [5], with the exception of the last one, which is from [16] (their terminology is different, however: our intensive / extensive substitutes the inclusive / exclusive adopted by the former work).

A particular may have only one volume, mass, density pattern, colour pattern. These are extensive determinations. However, a particular may have various colours and various densities. These are intensive determinations.

Extensive and intensive determinations are also distinguished by space and time, in that extensive determinations are extensive in at least one space-time dimension, and intensive determinations are intensive in at least one space-time dimension [5].

Table 1

	INTENSIVE IN TIME	EXTENSIVE IN TIME
INTENSIVE IN SPACE	colour, field strength, density, mixtures, tendencies, states, velocity, character traits	change (of velocity, of colour), point-formed motion
EXTENSIVE IN SPACE	volume, mass charge, voltage, shape, density and colour patterns	motion, function, action, work

The structure of types and kinds for particulars and determinations can be exemplified by the following diagrams relative to the string "red table weighing 10.5 kgs":

Expression	TABLE
Type	INDIVIDUAL
Kind	ARTIFACT

Expression	RED
Type	QUALITY: INTENSIVE
Kind	COLOR

Expression	10,5
Type	QUALITY: EXTENSIVE
Kind	WEIGHT

5 Names and values

We previously hinted at the difference between TABLE and TABLE’S WEIGHT. Let us now consider in general – that is to say, independently of its connection with the content TABLE – the content WEIGHT. This can be represented either as:

$$\frac{\text{weight: string}}{\text{WEIGHT}}$$

or

$$\frac{\text{10.5: floating}}{\text{WEIGHT}}$$

The two representation are evidently very different. The former provides one of the content’s names, the latter provides one of the content’s values.

Note that both representations are meaningful only if (1) their kinds and (2) their objectual referents are provided. The former may be specified as, say, naturalLanguage : English : CN for the string expression and as measurementUnit for the floating expression. The former may be specified as type : determination and kind : extensive for the string expression. In its turn, the floating expression would be characterized by type : ideal : mathematics and kind : $R \geq 0$.

6 Categorial units

We shall use the term ‘semiotic unit’ for the coordination between symbolic units (that is, expression and content units) and objects. As a shorthand way of representing the various distinctions, we shall adopt the following conventions:

- objects will be enclosed within square brackets, as in [Table]; signs will be written in small letters, as in table; contents will be written in capital letters, as in TABLE.
- types will be written as superscripts and kinds as subscripts.

Considering the first types and kinds for any unit, the ‘table’ exemplification can be represented as follows:

Object	$\text{TABLE}^{<C>}_{lD}$
Sign	$\text{table}^{S : lA}_{nL : E : CN}$
Content	$\text{[Table]}^{P : I}_{D : F : A}$

where: S = string, lA = latinAlphabet, nL = naturalLanguage, E = English, CN = countName, <C> = category, lD = labelledDescription, P = particular, I = individual, D = Dependent, F = Functional; A = Artifact. Note that ‘category’ is flanked by angled brackets. Their meaning will soon be apparent.

7 Assumption

The perspective outlined in the previous sections claims that a semiotic unit is codified by a three-fold structure: an expression, a content and an objectual (or ontological) component. Expressions and contents together give rise to the symbolic component of the semiotic unit. As we have seen, the three components are all further subdividable into types and kinds. Needless to say, their intertwined relationships become easier to elaborate if we find a uniform way to represent the various components.

We could formulate this requisite in a much stronger sense: analysis becomes easier as soon as we realise that all the previous components have a geometric nature. Since we cannot argue here in favour of this stronger thesis, we shall confine ourselves to its much weaker methodological translation, namely that analysis becomes easier if we find a *uniform* way to represent the various components of the semiotic unit. This is our major assumption.

8 Synopsis

Let us sum up the main distinctions considered so far. The following synopsis will give us an opportunity to introduce details not specified previously, viz.:

± two-branch tree; by default, subsequent values consider only the positively marked branch.

<> list of different values; the value indicated is one among many; subsequent values consider only sub-types of the indicated value.

The distinctions introduced so far form the following trees:

Expression unit
- Type : <string> : <latinAlphabet>
- Kind: ±naturalLanguage : <English> : <CN>

Content unit
- Type: <category>
- Kind: labelledDescription

Ontological unit
- Type: object: <substance> : <particular> : <individual>
- Kind: <dependent> : <functional> : <artifact>

The tree contains an undeclared shift. All the types of the content units have been unified under the label ‘category’. This move is precisely what is needed in order to obtain a uniform representation of the various components of the semiotic unit.

We stated earlier that labelled descriptions are the only kind of content unit, and that their types are mathematical structures. This is what was needed to start elaboration of a uniform perspective. Category theory in fact is that common framework in which many if not all mathematical structures can be uniformly analysed. Category theory therefore provides the answer to our major assumption.

To make our argument more plausible, we want to sketch the role of category theory in the construction of music knowledge spaces (for details, please refer to [11]). In this context of music semiotics, symbolic units are called denotators, whereas the semiotic units are called predicates. We first concentrate on the symbolic level of denotators. A denotator is defined by a quadruple (name, address, form, coordinates) with these specifications:

- name: this is a character string holding the name of the musical object;
- address: this is a module over an associative unitary ring R; for example R may be the ring of rational or real numbers, or the monoid algebra Z<ASCII> over the word monoid of ASCII symbols;
- form: this is a contravariant, set-valued functor on the category of modules, together with a recursive construction method (typically a universal construction, such as the limit, colimit or power object) which tells how the form is built from more elementary forms. Recursion may end up on representable functors on the category of modules, but circular definitions of forms are also admitted, and even important in the construction of musical concepts.
- coordinates: these data form a point of the evaluation of the present form at the given address. For example, if the form is a cartesian product of forms F_1, F_2,..., F_n, the coordinates is just a *n*-tuple of coordinates C_1, C_2,..., C_n for the respective *n* factor forms F_i.

In this setup, the category of modules parametrizes the ontological extent of symbolic units. For example it can be shown that a common piano note's address is the zero module, whereas the address of a contrapuntal interval is the module Z of integers. In fact, musicology views notes and intervals as 'living' in different ontological regions, and address modules take care of this subtlety. Note, however, that the address module does not fully cover ontology, it just symbolizes the ontological perspective of our knowledge space. More precisely, if the zero-addressed piano note is the symbolic unit targetting at a physical semiotic sound object, one shows that the Z-addressed contrapuntal interval targets at an ordered pair of piano notes, but it does so not only in this formal way, it really creates a new ontology: the interval object exists on its own right and is not only a crude couple of physical sounds.

This example also makes clear that knowledge deals with ordered access to information: Knowledge about music objects, as denotated by denotators, is operationalized on the basis of reference to components, and this reference is stamped by the order principles of universal construction of category theory.

9 Operators

From what has been said, it follows that other operators beyond the usual logical operators should be used. These are the 'geometric operators' of category theory. There are two arguments for this extension. First, the context of category theory strongly preconizes techniques and results from the categorial study of logic as it has been elaborated in the mathematical topos theory. Any text on category theory will provide all the necessary details (For an excellent presentation, see for instance [8]; for an application to cognitive problems see [9]).

Topos theory has succeeded in absorbing logic in categorial geometry, including models of classical and intuitionistic logic, geometric independence proofs of the continuum hypothesis, and a vast generalization of the concept of topology in algebraic geometry. So logic appears to be a very special case of topos-theoretic constructions.

The second argument is that our proposed geometric concept framework makes strong use of constructors which are genuinely geometric and not logic, for example fiber products or fiber sums. Recall that in computer science, data base management systems (DBMS, see [13] or [18]) also rely on entity types which follow the classification from topos theory; unfortunately, most DBMS theorists are not aware of this tight resemblence.

10 Ontological and quasi-ontological information

The analyses of particulars and determinations presented in §4 pertain to ontology. We have seen some aspects of the role ontology is supposed to play. To avoid misunderstandings, a further distinction must be drawn between properly ontological information and quasi-ontological information.

Ontology should be able to say that a certain thing is situated somewhere, or that an event has taken place at a certain moment. But it does not have to say these things using the Gregorian calendar or a particular system of coordinates. We choose a system of measurement for every magnitude, but which system is chosen is purely a matter of convention, and the relative module should be substitutable if for some reason it becomes necessary to use some other system of reference (with appropriate adjustments).

The same applies to many other aspects of design. Somewhere there will be a module in which the ontology is calibrated to the measurement systems employed, and to such other purely pragmatic aspects as the language of the user interface. Likewise, there must be a place in which naturals, connectives, some functions, and so on, are imported. But which particular version is used is not an explicitly ontological problem.

A further quasi-ontological category consists of what we can call 'signature'. This category furnishes information on who has made the categorization, where, when and how. Such information is not always relevant, but there are some contexts in which it is important: in medicine, for example, it is sometimes vital to know who has made a diagnosis. Aspects of this kind perform a role internally to a fully developed framework, but they are not directly ontological components [15].

11 Ontology vs epistemology

The previous pages have sketched one of the possible ways towards knowledge systems. We have seen that having systems rich not only in information but also in knowledge imply (or, at least, could imply) the thesis that our bases of data ought consider not only symbolic information (symbolic units) but also ontological information. We have also seen that properly ontological information should be carefully distinguished from quasi-ontological information. At this point a further step is still required, namely the proper distinction between ontology and epistemology.

Defining the precise tasks and characteristics of both ontology and epistemology is important if we are to avoid confusion between them; confusion that is often apparent in the literature. In fact, modelling (our knowledge about) objects and modelling our knowledge (about our knowledge about objects) are different tasks. The former is ontology, the latter is epistemology or theory of knowledge.

The difference can be evidenced by listing concepts of ontology and epistemology. Ontological concepts are: object, process, particular, individual, whole, part, event, property, quality, state, etc. Epistemological concepts are: belief, knowledge, uncertain knowledge, revision of knowledge, wrong knowledge, etc.

If ontology is the theory of the structures of objects, epistemology is the theory of the different kinds of knowledge and the ways in which it is used.

One objection that has been frequently raised runs as follows. In order to know what is represented by a certain representation, we must compare this representation with its object, and to conduct this comparison we must know the object independently of our representation of it. But the only way in which we can know an object is by means of its representations, and consequently we can never know an object independently of its representations. Thus we cannot know what a certain representation represents. This objection contains a subtle error: a representation represents an object, not the same representation. To know the representation we must have a representation of the representation. Therefore the premise that in order to compare a representation with its object we must know the object independently of its representation is false. To compare a representation with its object it is not necessary to know the object independently of its representation, which is simply impossible. It is instead necessary to know both the (representation of the) object and the (representation of the) representation of the object: we know the object through its representation and the representation through another representation which has as its object the previous representation. Our knowledge is first of all knowledge of objects: we know representations only in a second instance by transforming them into other representations ([4], 28-30).

A further much debated problem is whether domain knowledge can be represented independently of the way in which it is used in reasoning. In AI Clancey and Letsinger (see

[2]) claim that both domain knowledge and problem-solving knowledge can be reused (provided the two kinds of knowledge are represented separately in the knowledge base), while Bylander and Chandrasekaran (see [1]) argue in favour of the opposite thesis. The main argument developed by Bylander and Chandrasekaran is the so-called interaction problem, which can be summed up by saying that "representing knowledge for the purpose of solving some problem is strongly affected by the nature of the problem and the inference strategy to be applied to the problem".

We think that their diagnosis is correct, but the conclusions are wrong. The fact that there is a mutual or bilateral form of dependence between ontology and epistemology does not oblige us to conclude that we cannot represent their specific properties and characteristics separately. On the contrary, we should specify both what ontology can say about epistemology (a belief is a kind of object, it has parts and properties, etc.), and what epistemology can say about ontology (knowledge of the structure of objects is a kind of knowledge). This is a difficult task and mistakes are always possible, but there is no principled reason for denying its realizability, even if one understands why it is so easy to blur ontological and epistemological issues.

The ontological and epistemological perspectives interweave and condition each other in complex ways. They are not easily separable, amongst other things because they are complementary to each other. Ontology, in fact, is mainly bottom up, while epistemology is top down. In other words, these are two opposing perspectives. According to the first, in ontology knowledge is provided with a bottom-up foundation, whereas in epistemology a top-down synthesis reduces the object to the result of the application of cognitive schemata.

A further difference – similar but not identical to that between ontology and epistemology – is the difference between an ontological reading and an epistemological one.

Consider the sentence:

(i) Napoleon was the first emperor of France.

Its formal reading is:

(ii) Somebody was a something somewhere.

From a cognitive point of view, (i) may mean for instance that

(iii) The man portrayed by David in the likeness of a Roman Caesar was the first emperor of France [17].

(ii) clearly does not imply (iii), whereas (iii) does imply (ii). In general, it is always possible to develop many different cognitive readings of the same sentence. These various readings depend on the information that is implicitly or explicitly added. If we do not add new information, the reading (iii) above is unjustified because the sentence (i) does not entail the information that Napoleon was portrayed by David.

In general, (ii) (the purely formal reading) is too poor; it is general but it says too little. On the other hand, (iii) is too strong, it is not sufficiently general and it depends on added information. The real difficulty for the ontological reading is that it lies somewhere in between. It is more than the purely logical reading and it is less than the many different cognitive integrations.

The truly ontological viewpoint manifests the many facets of the object. It says not only that somebody was something somewhere, but that he performed an important institutional role (in fact the most important one) in some specific part of Europe, and it says that he was the first to perform that role. It says, moreover, that he was a human being and for this reason that he had a body and a mind, that he was alive, etc.

There is a myriad of information embedded in the sentence "Napoleon was the first emperor of France". The task of the ontological reading is to extract and organise this information without resorting to any external source of knowledge.

12 A provisional conclusion

Our conclusion can be stated as follows: passing from information to knowledge is a major step which involves the elaboration of a well structured theory of semiotic units. This step, in turn, requires both a well elaborated ontology and a proper setting of the mutual (or bilateral) dependencies between ontology and epistemology. From a formal point of view, the above seems to require formal tools at least as sophisticated as those provided by category theory.

References

[1] T. Bylander and B. Chandrasekaran, Generic Tasks in Knowledge-Based Reasoning: The Right Level of Abstraction for Knowledge Acquisition. In: B.R. Gaines and J.H. Boose (eds.) *Knowledge Acquisition for Knowledge-Based Systems.* Academic Press, London, 1988, vol. 1, pp. 65-77.

[2] W.J. Clancey and R. Letsinger, NEOMYCIN: Reconfiguring a Rule-Based Expert System for Application to Teaching. In: W.J. Clancey and E.H. Shortliffe (eds.) *Readings in Medical Artificial Intelligence: The First Decade.* Reading, Addison-Wesley, 1984, pp. 361-381.

[3] L. Hjelmslev, *Prolégomènes d'une théorie du language.* Éditions de minuit, Paris, 1968.

[4] R. Grossmann, *The Categorical Structure of the World*. Indiana University Press, Bloomington, 1983.

[5] I. Johansson, *Ontological Investigations*. Routledge, London, 1989.

[6] R.W. Langacker, *Foundations of Cognitive Grammar.* Vol I: *Theoretical Prerequisites*. Stanford University Press, Stanford, 1987.

[7] R.W. Langacker, *Concept, Image, and Symbol*. Mouton De Gruyter, Berlin-New York, 1990.

[8] S. Mac Lane and I. Moerdijk, *Sheaves in Geometry and Logic*, Springer, New York, 1992.

[9] J. Macnamara and G.E. Reyes (eds.), *The Logical Foundations of Cognition*. Oxford University Press, New York, 1994.

[10] G. Mazzola, Semiotic of music. In: R. Posner et al. (eds.), *A Handbook on the Sign-Theoretic Foundations of Nature and Culture*. Vol. III. de Gruyter, Berlin, 1999.

[11] G. Mazzola, *The Topos of Music*. Birkhäuser, Boston, 1999.

[12] J.G. Miller, *Living systems*. McGraw-Hill, New York, 1978.

[13] B. Plotkin, *Universal Algebra, Algebraic Logic, and Databases.* Kluwer, Dordrecht, 1994.

[14] R. Poli, Ontology for knowledge organization. In: R. Green (ed.), *Knowledge Organization and Change*. Frankfurt am Main, INDEKS Verlag, 1996, pp. 313-319.

[15] R. Poli, *Alwis. Ontology for Knowledge Engineers*, 1999 (forthcoming).

[16] S.S. Stevens, *Psychophysics. Introduction to its Perceptual, Neural, and Social Prospects*. Transaction books, New Brunswick (Usa) Oxford, 1986 (1st edition: Wiley, New York 1975).

[17] A.P. Ushenko, *The Field Theory of Meaning*. University of Michigan Press, Ann Arbor, 1958.

[18] G. Vossen, *Datenmodelle, Datenbanksprachen und Datenbank-Management-Systeme*. Addison-Wesley, Bonn, 1994.

[19] T.E. Wilkerson, *Natural kinds*. Avebury, Adelshot, 1995.