Claudio Gnoli (University of Pavia)

Workshop on Levels of reality as a KO paradigm
Levels, types, facets: three structural principles for KO

Abstract

Three major principles in the structure of knowledge organization systems are identified and discussed: hierarchical trees of types, sets of facets conforming to general categories, and series of levels of reality. Each principle can be present at various degrees in different systems. The three principles can interact between them in various ways, depending on priority choices in system construction. Examples are reviewed of different priority options adopted in classifications and thesauri, both special and general, and their effects are critically examined. It is found that levels of reality, although less often explicitly acknowledged than other principles, contribute an important tool for knowledge organization.

1: Introduction

In order to discuss the theme of this workshop, I mean to make here some observations on how levels are used as a structuring principle in knowledge organization systems (KOS), and how they interact with other structuring principles.

In knowledge organization (KO) literature, two major principles are widely acknowledged to describe the structure of KOSs: hierarchical trees of types, and combinations of facets. In the following, these principles will be reviewed, and their relationships will be examined with the additional principle of levels, which also plays a relevant role in the structure of many KOSs, though often not explicitly. We will then consider how these three principles can interact between each other, working as sub-structures that concur to determine the macro-structure of a KOS.

2: The three principles

In this section I will briefly describe the three principles considered here as major structuring elements of KOSs: types, facets, and levels. Other principles can also be identified, like disciplinary aspect (Slavić 2007) or knowledge dimensions (Hjørland & Hartel 2003). However, types, facets, and levels provide three patterns general enough to account for many basic features of KO structures, without at the same time producing a proliferation of models, that would make this initial analysis exceedingly complex.

Each principle can also be described as a kind of relationship between concepts. Indeed, all KO structures can ultimately be reduced to concepts, that can be treated as entities, and relationships between them (Green 2008). Relationships are basic structural elements especially in such KOSs as thesauri, ontologies, and conceptual maps.

2.1: Types

Tree-like division of the objects of knowledge into types is the most classical principle of organization. This is what immediately occurs to mind when speaking about taxonomies or classifications. A class (genus) can be divided into two or more subclasses (differences or species) by application of a certain characteristic of division. The process can be repeated at each stage, thus producing a tree-like pattern. This procedure is also known as “Porphyrian tree” and has long been taught in classical logic. Successive branchings represent concepts more and more specific, that is, laying
at increasing degrees of depth (sometimes also called “levels”, which however generates a terminological confusion between levels and types).

Decimal classifications represent such depth by an increasing number of digits or letters. Types are also a basic structuring principle in thesauri, where they are expressed by the BT-NT relationship (Beghtol 2000). The relationship between a class and its types is usually called “hierarchical”, or “subsumption” relationship. It can be recorded in a form of the kind Vertebrate isA Animal. This, however, does not hold for whole-part relationships, as a tail is not an animal. Indeed, some KOSs treat parts not as types but as facets.

2.2: Facets
Facets are attributes of knowledge objects, that are represented as complementary terms or as notational suffixes of a basic class. The notion of facet and its systematic representation in KOSs has been introduced by S.R. Ranganathan (Beghtol 2008), although in a broader sense common auxiliaries of time, space or form, already used before Ranganathan, can be considered as facets as well. In computer science and information architecture, facets are described as the possible “values” of some “attribute” (Priss 2008), often corresponding to one field in a database, while in some ontologies they are called “slots”.

The attribute represented in a facet can be of any kind: we can speak generically of an associative relationship, the RT relationship in thesauri. It can be recorded as Animal hasAttribute Tabby. However, in faceted KOSs the kind of attribute is expressed explicitly, by saying e.g. that animals have colour, weight, age, diet habit as their facets. Each domain has a characteristic set of facets, that can be identified empirically by analyzing large enough samples of literature in that domain, then a helpful citation order can be established between its facets (Vickery 1960).

It has been found useful that the facets specific of each domain be referred to a pattern of general categories valid for any domain, like processes, actions, agents, materials, or parts. This allows to follow a standard citation order of categories across the different domains covered by a KOS, contributing to its internal consistence. Speaking of facets or general categories is thus equivalent for the purposes of this paper, as they play the same structural function. General categories are very interesting from the philosophical viewpoint, as it can be discussed whether they have an ontological or an epistemological foundation, whether they really are generally valid across all domains or not, etc.

2.3: Levels
Levels seem to be the least often acknowledged among the three principles under discussion. Levels of reality consist in a set of features in the known world that can be arranged, at least intuitively, in a natural series, such that the existence of each level depends on the previous ones (Poli 2001). Frequently cited levels include matter, life, mind, and society, although many authors have proposed their different versions of the list, with variable degrees of detail (references in Gnoli & Hong 2005).

We can recognize e.g. life as a level higher than matter as, according to the present state of human knowledge, life can exist only in presence of matter, all living entities being made with material elements. At the same time, life is something more than
matter, as it shows characteristics and properties that are not shown by matter, such as metabolism, homeostasis, and reproduction. Thus an accurate description of living entities requires the use of notions that are not present in the theory of matter, but are peculiar of the theory of life (Poli 2010). The relationship between different levels can be described as one of existential dependence (Gnoli et al. 2007). It is a central claim of the theory of levels that this relationship is asymmetrical: while life depends on matter, matter does not depend on life, as material entities also occur in many contexts where life has never appeared, like e.g. the body of stars. In a KOS, this could be expressed as Life dependsOn Matter.

In Nicolai Hartmann’s ontology, two kinds of dependence between levels are identified: one of “overforming”, holding between minor levels (layers) which are materially made of elements of the lower ones; and another of “building-above”, holding between major levels (strata) which, though depending on the lower ones, have not them as their material constituents (Hartmann 1942; Poli 2001; Scognamiglio 2010). This second kind of dependence appears to be the most mysterious, although it has been suggested it always corresponds to the appearance of some mechanism of memory (Jacob 1970; Gnoli 2008).

A sequence of levels is sometimes referred to as a “hierarchy of levels”: e.g. living entities are said to have a “hierarchical” structure consisting of the levels of biomolecules, cells, tissues, organisms, and populations. Unfortunately, this phrase again introduces confusion with the notion of hierarchy of types; for this reason, it is advisable to avoid it in KO contexts. Another classical image of levels, coming from medieval Europe, is that of a “great chain of being”, including rocks, plants, animals, men, spirit, and God (Lovejoy 1933). Levels correspond to links in the chain, or to stair-steps in the “scala naturae” (nature ladder) (Barsanti 1992). These metaphors reflect the dependence of higher levels on lower ones, representing it as the action of gravity.

In modern science, the series of levels is seen mostly as an evolutionary one: each new level is thought to have appeared at some point in the history of the universe, by a process of emergence. This idea has been supported by the philosophical movement of emergent evolutionism (Blitz 1992). Levels and their branchings can thus be interpreted as an evolutionary tree, suggesting a phylogenetic model for KO (Gnoli 2006).

The theory of levels was explicitly acknowledged as a useful principle for KO by Douglas Foskett and other members of the Classification Research Group, who were looking for a criterion to arrange concepts in a new general classification system, that should not have been based on canonical disciplines anymore (Foskett 1961; 1978; Spiteri 1995; Gnoli & Poli 2004). The CRG received it in its versions by Joseph Needham and James Feibleman, both speaking of “integrative levels”. CRG ideas have been reconsidered recently in the Integrative Levels Classification (ILC) research project (Gnoli 2008). Dahlberg (1982) also bases the main classes of her Information Coding Classification (ICC) on a series of ten levels; similar attempts have been done by Wählin (1963) for his Universal System, by Scheele (1977) for his Universal Facet Classification, and by Shpackov (1992) for his Universal Classification. More generally, in many general classifications the sequence of main classes corresponds to a series of levels, although this is not always declared explicitly (Gnoli 2006; Dousa 2010).

The usual application of levels in KO is to arrange classes within an array. This is done especially for the main classes of general KOSs: this can thus be seen as the
identification of levels in reality, rather than the division of the universe into an arbitrary set of main types. It is disputed whether the arrangement of levels in a linear sequence is natural or partially imposed for practical reasons, that is if they form a total or a partial order. The principle can work also within more specific arrays: Foskett suggested the cases of language units (morphemes, words, phrases, sentences, discourses) and of social units (families, groups, cities, nations, international organizations).

3: The problem of priority
After observing that each of the three principles is widely used in KOSs, we have now to consider how they are combined. There are several logical possibilities for this. Each principle can be either present or absent (we saw that facets and levels often are present only implicitly). Extreme KOS varieties using only one principle are not very useful, and in most cases some combination of more than one is used. This implies a need to ascertain at what step of construction the principle is used. The overall structure must have a dominant principle, while others can be used for more detailed articulation. In other words, principles have to be applied in some order. In this sense, many possibilities have been explored by existing KOSs. We will review some notable examples of them in next section. In order to reduce the combinations to a manageable number, we will consider only ordered couples of a primary and a secondary principle. For each couple, the missing third principle can sometime act as a tertiary structure for further refinement into more specific concepts.

4: Priority options
4.1: By types, then levels
In classical enumerative classifications, like Dewey Decimal Classification, Cutter’s Expansive Classification, Library of Congress Classification, and Universal Decimal Classification, the universe of knowledge subjects is divided primarily into main classes of types. These are determined usually on the basis of epistemological characteristics, like the different knowledge faculties (reason, imagination, memory) which according to Bacon produce the different disciplines. Main classes are further divided into sub-types, sub-sub-types, etc. However, levels appear as a principle to sort some arrays of types at a given degree of specificity, e.g. within the type of pure sciences, these are listed in the level sequence of mathematics, physics, chemistry, biochemistry, biology. Levels here work as a “principle for helpful sequence” in the forms of “later-in-time”, “later-in evolution”, or “increasing complexity” (Ranganathan 1967, part F).

4.2: By levels, then types
When the approach to the division of knowledge is ontological, that is referring to the known objects rather than the faculties by which their are known, levels become an obvious way to order these objects. As we saw above, this was done by classificationists like Brown and Bliss. Their systems list main classes themselves in an order of “gradation by speciality”, basically corresponding to a sequence of levels of increasing organization. Within each level, types of objects are identified and listed, like painting, sculpture, and music within the art level. The types by which a level is divided are sometimes also listed in an evolutionary sequence: the level of organisms can be
divided into the kinds of protista, fungi, plants, and animals, which are again in an evolutionary sequence. When considering a new class, like Algae or Men, the classificationist has to decide whether they are a subclass of an existing one, or are different enough to create a new class of the same degree, that is to acknowledge one more “grade” in evolution (Huxley 1958; Gnoli 2006). In this phylogenetic perspective, types tend to coincide with levels. Indeed, the types of organisms are also different levels reached by the historical process of evolution; similarly the class of religions can be divided into the various historical religions, which have chronological and phylogenetic relationships between them (Broughton 2000).

4.3: By types, then facets
Facets are promoted in the second place in explicitly “faceted” systems like Colon Classification. The primary division is still in types, but each type is well structured by the application of facet analysis within it. However, the facets of one type remain autonomous from those of another type, as types prevail on facets. Each type is thus a quite isolated domain. Levels only appear as a tertiary principle to sort the subclasses of a type, or the foci of a facet.

4.4: By facets, then type
In special classifications and thesauri, the domain corresponding to one type in a general KOS is assumed as the default context. Thus the structure described in the previous paragraph can be cut on its typological top, leaving facets to become the primary principle. Systems like the London Education Classification, and the Art and Architecture Thesaurus are divided primarily by facets, and these are further divided into types. This structure is also adopted in general verbal KOSs, such as the Nuovo Soggettario (NS) developed at the Central National Library of Florence (BNCF 2006). In these cases, facets are not specific of any given domain, but take the form of general categories, like Things, Properties, Activities, Time. We have thus a tree hierarchy within the Things facet, another within the Properties facet, etc. Cats will belong to the tree of things, while Tabby will belong to the tree of properties.

4.5: By levels, then facets
Some KOSs primarily structured by levels also adopt facet analysis. This is the case of BC2, having main classes corresponding to levels, which are then divided by facets. This means that each level has its own characteristic facets: organisms have the facets of physiological process and habitat, while artifacts have the facets of raw material and purpose. Facets even contribute to the definition of the level itself as a “categorically homogeneous region”: the things that have physiological processes are organisms (Gnoli 2008; Poli & Obrst 2009; Poli 2010).

After being divided them by levels and facets, concepts are usually divided further by types. An exception to this rule is ICC: it is divided primarily into ten levels, then these are divided by ten categories (the “Systematifier”), then again by the same categories, and so on, thus identifying more and more specific knowledge fields (Dahlberg 2008). Types never appear in this process, although Dahlberg claims that more specific types of knowledge objects could be identified in the future development of the scheme.
The interplay between levels and facets/categories was considered by the CRG during the NATO research for a new general classification scheme. The two logical alternatives, giving priority either to facets or to levels, were acknowledged. For example, Ranganathan’s Energy (Activity) category could now work either as a primary principle of division, like in faceted thesauri, or as a principle to divide single levels of Personalities (Things):

“Mr Foskett suggested that Energy isolates for a general classification should be sought in the first place in the literature dealing with a particular “Thing”. It might be found that a given activity was associated either with a single “Thing”, or a group, or with all “Things” at the same integrative level, or even with all “Things” at all levels. Only the last should be regarded as isolates in the category of Activities, the remainder being dealt with in a manner analogous to the treatment of differential facets in the Colon Classification. Each such differential Activity might be placed in linear order before the single “Thing”, group or level of “Things” with which it was exclusively associated.” (CRG 1961, 163)

Foskett’s preference offers the advantage that levels can thus be listed only once, with all activities, properties etc. which are specific of that level being filed under it in the schedules. Physiological processes will be listed as a facet of organisms, as their definition is interdependent with that of organisms. Only general activities, like change or appearance, would be listed in a general class of activities before all levels. Both the main account of the NATO project (Austin 1969a-b) and the ILC draft scheme indeed conform to this idea.

4.6: By facets, then levels

In Austin’s project, general attributes that can be applied to any level are listed before all levels, and divided into Relative terms, Positional terms, Properties, and Activities. For these concepts, facets are thus the primary principle of division. This architecture is adopted in many thesauri, where facets are the primary structuring principle, before types (BT-NT). This produces the result that physiological processes are listed under activities, while organisms are listed under things, so that the notational symbols of physiological processes and of organisms have no relationship. In this way, a combined notation to mean “physiological processes of organisms” would be lengthy and have a redundant meaning, as physiological processes would refer to organisms both by definition and by combination.

Since the 1970s, Austin devoted himself to the development of PRECIS, a verbal general KOS where levels were not used as a structuring principle anymore. However a successor of PRECIS, Nuovo Soggettario (NS), is now beginning to consider again levels as an additional structuring principle besides types and facets. Like many thesauri, NS is primarily structured into facets (of which the most general are called “categories”): Things, Actions, Agents, and Time (Cheti & Paradisi 2008). These are then divided into types (“facets”, “subfacets”, and their narrower terms).

However, due to the uncommon case of NS being a general thesaurus, its developers have faced themselves with some problems concerning the definition and listing of some concepts. Generic action terms, such as management, have to be disambiguated in order to distinguish between the management of a company from the management of land resources in a country. A first solution has been the addition of scope qualifiers in the form of management <economics>. This, however, would imply that disciplines have a definitional role for the whole system, while this is not intended in NS original design, as concepts essentially represent phenomena (disciplines are rather represented...
Another solution can be offered by considering levels as an additional structuring principle. This opportunity is now being considered, in the light of the original CRG tradition from which both NS and ILC descend. In the logic of integrative levels, the concept of management has a place of unique definition laying at the level of social organizations. Ideally this concept should be defined at the organizations level, and combined with concepts at any other level only in the syntactical stage. However, as the NS structure takes facets as its primary principle, it is the facet of Actions that will be divided into material actions, organic actions, mental actions, and social actions. In the same way, Things could be divided into material things, organic things, mental things, and social things, etc. This will produce a system structured by facets, then levels, then types.

5: Concluding remarks
We have seen how levels, types, and facets can interact and be combined as three major structuring principles in KOSs. Of these, types are the most classical principle. Facets were studied and applied later, but can now be seen as another classical principle; indeed, they are also widely acknowledged in recent literature on information architecture and digital content management. Levels were acknowledged even later in an explicit way, by the specific studies of the CRG. However, their roots can be traced in ancient notions like that of scala naturae.

Furthermore, it has been shown how levels often act as an implicit principle. An explicit analysis of their role, and of the different opportunities of their interaction with the other principles, is believed to be a useful contribution to the development of a more complete theory of KO. Recently spreading KOSs like general thesauri, taxonomies, and ontologies show a poor awareness of their possible role. Still some cases, like the Nuovo Soggettario discussed above, or the General Formal Ontology (Herre et al. 2006), show that levels can fruitfully contribute an additional tool for the development of KOSs capable of reflecting the many aspects of reality in more accurate and effective ways.

Acknowledgments
Many thanks to Roberto Poli, Carlo Scognamiglio and Alberto Cheti for useful comments to this paper and, more in general, relevant discussions about levels of reality and the ontological principles of KO.

References
Austin D., 1969a, The theory of integrative levels reconsidered as the basis of a general classification, in Classification Research Group, Classification and information control, Library association, London, p. 81-95.
Barsanti G., 1992, La scala, la mappa, l’albero: immagini e classificazioni della natura fra Sei e Ottocento, Sansoni, Firenze.
Beghtol C., 2008, From the universe of knowledge to the universe of concepts: the structural revolution in classification for information retrieval, Axiomathes, 18, n. 2, p. 131-144.


CRG, 1961, Classification research group bulletin n. 6, Journal of documentation, 17, n. 3, p. 156-172.


Hjørland B., Hartel J., 2003, Afterward: ontological, epistemological and sociological dimensions of domains, Knowledge organization, 30, n. 3-4, p. 239-245.


Scognamiglio C., 2010, “Strata” and top categories for an ontologically oriented classification, in *this volume*.