Roberto Poli
Department of Sociology and Social Research, Trento and Mitteleuropa Foundation, Bolzano, Italy

Framing Information

Abstract: The distinction between semiotic, semantic and ontological classifications is introduced. A few examples of semantic and ontological categories are then provided and discussed. The thesis is defended that semantic categories depend on ontological categories.

1. Introduction

One of the most striking features of recent developments in science and technology is the over-abundance of information. 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 information. On the other hand, 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 events and objects of our experience are classified in many different ways: some forms of classification depend on the way in which events and objects are perceived (seen, felt, heard, smelt, tasted, together with the interactions among these perceptions); others depend on the way in which reality is conceptualized (e.g., in terms of its salient features); others depend on the ways in which perceptive and cognitive analyses are codified in linguistic expressions and structures; and yet others depend on general patterns or universals wholly intrinsic to the events and objects of the world. I distinguish three basic types of classification: the semiotic, the semantic, and the ontological.

Semiotic classifications distinguish between *codes* (linguistic, musical, pictorial, etc) and set the modeling environment for their proper analysis. This means that the usual series of types (string, floating, integer and boolean) should be extended 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 semiotic dimension therefore analyses the features of the signs used to convey information. Its main, but by no means only, sub-section is the language analysis module (semiotic aspects are further analyzed by Poli and Mazzola 2000).

Semantic classifications analyse forms of contextual and functional dependence. Moreover, they collect and seek to coordinate both *de facto* and formally established standards. (I have reluctantly decided to resort to the overabused "semantic" owing to my inability to find a better term.)

Ontological classifications analyse the deepest internal (structural) features of the items under analysis.

The paper is divided in two sections, plus this introduction and a few words of conclusion. Section 2 presents a few data on semantic classifications and Section 3 will analyse ontological classifications.

2. Semantic classifications

Besides semiotic distinctions, we also need a large quantity of semantic categories if we are to describe and classify items. The easiest point of departure for semantic categorization is the *unity* of the item. One then proceeds by classifying the item with respect to related criteria of analysis. Before I continue, note that the just mentioned point of departure for semantic classification depends on the ontological category of *whole*: an item has unity (i.e., is one) if and only if it is a whole. It is therefore evident that ontological categories underlie semantic categories. I shall return to this problem in Section 3.

Subsequent steps concern the categorization of (application of categories to) the item(s) under analysis. Here I shall concentrate on the widely accepted strategy which relies on the distinction between natural kinds and dependent kinds. In general, there is broad agreement that electron, proton, neutron, narcissus, chimpanzee, stickleback, carbon, gold and water are natural kinds, whereas table, nation, banknote, rubbish, cliff, perennial and bush are not. It is likewise generally agreed that, if there are natural kinds, they fall into at least two groups. There are *kinds of stuff*, such as carbon, gold, water, cellulose, and there are *kinds of object*, such as tiger, chimpanzee, stickleback, narcissus (Wilkerson 1995). As before, *stuff* and *object* are ontological categories. Besides the trivial cases of the categories stuff and object, there is nothing to prevent acceptance of other ontological categories, like those of process and group (and perhaps even further less acknowledged categories: for details see Poli 2001).

Dependent kinds may be further subdivided into *functional kinds*, such as table and banknote, and *contextual kinds*, such as cliff and bush.

Crystal-clear cases aside, it may be difficult to properly distinguish cases of natural, functional and contextual items. The following may be one possible (partial) way out of the problem for the natural/functional opposition. Let us assume that functional kinds are linked to *technologies*, that is to say, to practices employed to name, recognize, use, produce and modify items of various types. The argument that I wish to develop is that the relation which connects natural and functional kinds is of the same nature as the relation between sciences and technologies. For this reason, it is advisable to dwell for a moment on the difference between sciences and technologies. The key reason for linking natural kinds to the sciences is that the latter, however much they are obviously interconnected, represent irreducibly different points of view. “An excellent reason for taking biology seriously is that the biological properties of things obviously depend directly on their physical and chemical properties. But the explanatory apparatus of biology cannot in practice be reduced to the explanatory apparatus of physics or chemistry or both” (Wilkerson 1995, 39).

In other words, the explanatory apparatus of sciences is emergent. Any science has its own ‘window’ on the world, selecting (in its own way) only those objects that are at the ‘right’ level of magnitude, energy and complexity. On the other hand, the explanatory apparatus of technologies is not emergent. “It is, as it were, … constantly being reduced to, or connected with, or supplanted by, the explanatory apparatus of some discipline characteristically concerned with entities of some lower level, notably physics, chemistry and biology” (Wilkerson 1995, 40).

Therefore, we may say that items described by sciences are items of a certain natural kind, whereas items described by technologies are items of a functional kind. Obviously, there is no clear-cut distinction between sciences and technologies: science is involved in numerous technologies, and technologies use the most disparate of sciences. Even if the proposal just put forward may not be able to settle all the possible cases, it nevertheless provides some guidelines.

The other remaining case concerns contextual kinds. The latter can be well represented by the following examples:

> Gardeners talk cheerfully of seedlings, saplings, trees, shrubs, bushes, climbers, perennials, annuals, pot plants, and so on, but none of these terms pick out a real essence; none are likely to appear in reports of serious scientific investigation; and none refer to a kind determined by an intrinsic property. One and the same plant will grow as a tree under one set of conditions and as a shrub under others (e.g. many *Eucalyptus* and *Acer* species). One and the same plant will be an annual or pot plant in a temperate European climate and a shrub in a hot African climate (e.g. *Pelargonium* species). One and the same plant is a shrub in western Ireland and a hardy perennial in Nottingham (e.g. *Fuchsia magellanica*). None of those terms pick out an intrinsic property and none of them correspond, even approximately, to any botanical classification ... Yet none of the terms has any connection with convention, artifice or culture (Wilkerson 1995, 37).

The same point can be made about geographical and meteorological kinds:

> Geographers talk of beaches, cliffs, mountains, valleys, seas and volcanoes. Meteorologists talk of depressions, anti-cyclones, winds, thunderstorms, clouds and hurricanes. But the terms do not pick out things with real essences, they do not figure in scientific generalisations and they do not pick out any relevant intrinsic properties. One and the same lump of material will count as a mountain in one environment, as a valley floor in another, and as part of the sea bed in yet another (Wilkerson 1995, 37-8).

The main feature of contextual kinds is that they contain an essential link with their enviroment. If you change the environment, contextual items may present different (even radically different) features.

Further developments concern hybrids between natural and dependent categories or between functional and contextual ones (for lack of space I cannot dwell any further on these: for details see Wilkerson 1995 and Poli 2001).

Before concluding this section, the problem of the many dependence relations between kinds should be briefly considered. Consider for instance the connection between the natural kind *wood* and the functional kind *table*. From the point of view of categorial analysis, what are the features of the complex whole *table made of wood*?

Put otherwise, what information is conveyed by its natural component (wood) and what information is conveyed by its functional component (table)? It takes a little thought to see that in this regard we are not too distant from Aristotle. If we know that the object over there is a table, we can make safe predictions about its behaviour in certain circumstances. In fact, we know the likely outcome of putting the kitchen table on a bonfire (Wilkerson 1995, 34). The point to the stressed is that, in making our predictions, we are exploiting the fact that *every item belongs to at least one natural kind or is dependent on an item belonging to at least one natural kind*.

In its turn, the theory of dependence referred to here requires a theory of levels of reality (see section 3 below). Let us suppose that we have developed one. Thus the general thesis becomes: *for each of their ontological levels*, all items belong to at least one natural kind and to one or more dependent kinds (functional or contextual). For any level, the natural kinds of every item are connected with the causal links of the item (as the above example of the table on the bonfire indicates).

This section has presented a few of the intricacies lurking behind a fragment of the field of semantic classifications, the one that relies on the distinction between natural and dependent kinds. The two main results of such analysis is that (1) semantic categories rely on ontological categories, and (2) items having a pluristratified nature are such that at least some of the subitems pertaining to one of their levels belong to a natural kind. (Other forms of semantic classification besides the opposition between natural kinds and dependent kinds are discussed in Poli 2000 and 2001.)

3. Ontological Classifications

Of the various types of classification, the ontological one is the most difficult, for a number of reasons (some of which have been analysed by Poli 2001 and 2002). In this section I shall try to give a very simplified idea of the highly complex structure of ontology.

The first step is to distinguish between what we are talking about and its determinations. Resorting to traditional terminology, I shall address the first topic as the problem of "substance", and the second one as the problem of "determinations". For the time being, I shall only consider the problem of substance.

My basic tenet is that the theory of substance comprises at least three sub-theories: the theory of particulars, the theory of levels of reality, and the theory of wholes and their parts (all of them alluded to in the previous section). Most traditional and contemporary theories of substance fail precisely because they lack one or more of the above sub-theories.

Let us continue to use "item" as the most generic descriptive term. Subsequent distinctions should consider at least objects, processes, stuffs and groups. Each of these has its ontological features and deserves its own theory. Higher-order items are items composed of other items (groups of processes, as for parallel computing, etc). (Poli 2001 distinguishes eight different types of basic item.)

As to the problem of levels, we may distinguish at least three ontological strata of the real world: the material, the psychological and the social. Specific forms of categorial and existential dependence exist among these strata. For example, a psychological item or event requires an animate physical object as its existential bearer. Should there be no person (and should there be no body of some such person), then neither will there be the correlative psychological states.

A relationship of matter and form holds among many items. In these cases, matter and form are correlative categories, so that any form may be the matter of a higher form, and any matter may be the form of a lower matter. The hierarchy thus constituted is a progressive overforming of matter and form. The nature of the physical world is clearly governed by this embedding principle: the atom is the matter of the molecule, but it is already an entity endowed with form; the molecule is the matter of the cell; the cell is the matter of the multi-cellular organism; and so on.

However, not all the dependences that structure the world are of a matter/form type. When one moves from the organic to the mental plane, one finds a dependence relation that is not reducible to the matter/form relation. One cannot say, in fact, that atoms or cells or organisms are the matter of the mind. Organic reality takes atoms and molecules and assembles them into a new form, consciousness, which is nevertheless not made up of organic forms. In the passage from the material to the mental there arises a *new* series of forms whereby corporeal life with its forms and processes no longer functions as matter. The organic levels are mirrored in psychic life: they influence it, they follow close upon it, but *they are not part of it*. In effect, the life of the mind does not comprise organic processes, nor does it use them as its building blocks, even though it is supported by them and is influenced by them.

One finds another break in the social stratum. In both these cases the dependence relationship is no longer of matter/form type but becomes one of a completely different kind: a bearer/borne relationship. In these cases, the substratum of the higher level is not the matter of the lower level (Hartmann 1952, 68-69).

Belonging to the social stratum are all phenomena of communication, and therefore the complex of social phenomena and customs, economic and legal realities, history, language, science, technology and the body of knowledge of every epoch, and morals.

Description of the strata and levels of reality intersects with description of the items of which it is composed. We humans participate in all three strata (although we do not exhaust the multiplicity of any of them). We have a material (organic) base, we have a mind, and we are simultaneously social beings. But our material base is one of the many material bases offered by the natural world; just as our mind is only one possible mind, and our participation in the social world is never such that we can absorb it in all its aspects (Poli 1998, 2001).

The third dimension of the problem of the substance comprises the theory of wholes and their parts. Without entering into detail, three kinds of wholes at least may be distinguished: aggregates, wholes in the proper sense and systems. Aggregates consist of proximate parts. Wholes in the proper sense comprise parts "which go together". Systems require a dynamic exchange between the whole and its parts.

Unity by solidarity is stronger than unity by proximity. This means that only some items that are aggregates are also (integral) wholes, and that some of the latter may be systems. Needless to say, the most difficult task is furnishing an adequate characterization of the dynamic components of systems. For the moment I merely point out that it is possible to determine various forms of dynamic unity, ranging from those that obtain in material systems of a physical nature to those that obtain in systems which, like living and social systems, are able to produce the elements of which they are composed.

The above classification can be further clarified by adding that aggregates are characterized by relations among their parts. Wholes require both part-part relations and parts-whole relations, whereas a proper characterization of systems requires information on three kinds of relations: part-part relations, part-whole relations and whole-part relations (Poli 2001, ch. 7).

Even if the above is only a very rough sketch of the basic structure of an ontology, it may nevertheless provide the reader with initial understanding of its main tenets. The point is relevant because, for some time, the term and idea of ontology have begun to enjoy currency in various sectors of the information processing community (and particularly in groups working in the fields of (i) knowledge representation, (ii) databases, (iii) natural language processing, and (iv) automatic translation). In short, those who most frequently talk about ontology are researchers in the acquisition, integration, sharing and re-utilization of knowledge (cf. Guarino and Poli 1995, Poli 2001).

Ontology comes into play as a viable strategy with which, for example, to construct robust domain models. An ontologically grounded knowledge of the objects of the domain should make their codification simpler, more transparent and more natural. Indeed, ontology can give greater robustness to models by furnishing criteria and categories with which to organize and construct them; and it is also able to provide contexts in which different models can be embedded and recategorized to acquire greater reciprocal transparency (Poli 1996, Poli and Mazzola 2000, Poli 2001).

Traditional philosophical ontology (like the one sketched above) and its new understanding (call it ontology as a technology) are still very distant from each other. This comes to no surprise: ontology as a technology is still in its early stages of development. At the moment, the artificial intelligence research community seems to have reached broad agreement only on the need for formal standards. This is indubitably an important step forward, but it is one which we can call authentically ontological only because of its improper extension of the concept of ontology (understood in its strict sense as categorical analysis; on this more usual sense of ontology see Poli 1992, Poli and Simons 1996).

An example may be of help. It is well known that important standards for software construction have become established in recent years. Suffice it to mention the Standard

Template Library (STL) for C++. This is certainly a positive development, but it is one that concerns solely formal components, not ontological ones. We will be able to talk of a similar development in ontology when we have a Standard Template Library for ontological patterns, like PROCESS, THING, EVENT, PART and WHOLE (Poli 2001).

We have seen that ontology has a complex conceptual structure. Without any awareness of its complexity, it is not possible to devise the most appropriate methods of analysis. On the other hand, it is the world that is complex, and there is no reason why it should be amenable to hypersimplified codification.

4. Conclusion

The paper has defended the theses that there are at least three main different types of classifications (semiotic, semantic and ontological), and that at least some semantic classifications require and rely on underlying ontological categories. Unfortunately, most contemporary scholars seem unaware of the challenging complexity of the ontological framework and this may undermine the value of even major projects.

References

Guarino, N. and R. Poli (eds.). (1995). *The Role of Formal Ontology in the Information Technology* (International Journal of Human-computer Studies), special issue, **43**, pp. 623-965.

Hartmann, N. (1952). *The new ways of ontology*. Chicago (Engl. Transl. of “Neue Wege der Ontologie”).

Poli, R. (1992). *Ontologia formale*. Genova, Marietti.

Poli, R. (1996). Ontology for Knowledge Organization. In R. Green (ed.), *Knowledge Organization and Change*. Frankfurt /Main, INDEKS Verlag, pp. 313-319.

Poli, R. (2000). Elementi per una piattaforma di categorizzazione semantico-ontologica. *Axiomathes*, **11**, pp. 103-136

Poli,R. (2001). *Alwis. Ontology for Knowledge Engineers.* Utrecht.

Poli, R., (2002). Descriptive, Formal, and Formalized Ontologies. In D. Fisette (ed.), *Edmund Husserl’s Logical Investigations 1901-2001: Origins and Posterity of Phenomenology*, Kluwer (in press).

Poli, R. and P. Simons (eds.). (1996). *Formal Ontology.* Dordrecht, Kluwer.

Poli, R. and G. Mazzola (2000). 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, pp. 1-11.

Wilkerson, T.E. (1995). *Natural Kinds.* Avebury, Adelshot.