

Roberto Poli

Department of Sociology and Social Research, Trento, Italy

Ontology for Knowledge Organization

Abstract: The first main thesis of the paper is that an ontology is not a catalogue of the world, a taxonomy, or a terminology. If anything, an ontology is the general framework within which catalogues, taxonomies, and terminologies may be given suitable organization. The second main thesis is that reality is organized into diverse levels and there are sophisticated dependences among these levels and within them.

1. Introduction

In the literature, the term 'ontology' presents a variety of meanings and is used in many different ways. To verify this observation one need merely consult any dictionary of philosophy. As well as the variety of meanings entrenched in the philosophical tradition, new variants have been recently forthcoming from the database community and from certain sectors of natural language analysis. For example, Sowa states that, "The first step in designing a database, a knowledge base, or an object-oriented system is to select an appropriate collection of ontological categories" ([Sowa 1995]. For discussion of the relationships between AI and ontology see [Guarino and Poli 1995]. More in general, on ontology and with particular regard to formal ontology, see [Burckhard and Smith 1991], [Poli 1992], [Poli and Simons 1996]). Apart from their multiplicity, many of these diverse meanings are accompanied by hidden assumptions and research strategies which are only occasionally made explicit. With regard to ontology, moreover, for some curious reason suggestions that the development of ontological analysis might be advisable often provoke extreme and contrasting reactions. While some regard ontology as a panacea for an extremely wide range of problems, others adopt entirely the opposite view and deny that ontological inquiry has any sense at all. If on the one hand there is often a desire to 'capture' the world in a definitive set of categories, on the other there is the opinion that the endeavour is impossible and indeed nonsensical. In fact, both points of view illegitimately overstretch the argument. The proposal advanced below seeks to avoid both baseless enthusiasm and deligitimating rejection. For simplicity's sake, I shall proceed by advancing a number of theses.

2. A number of general theses on ontology

THESIS 1. An ontology is not a catalogue of the world, a taxonomy, a terminology or a list of objects, things or whatever else. If anything, an ontology is the general framework (= structure) within which catalogues, taxonomies, terminologies may be given suitable organization. This means that somewhere a boundary must be drawn between ontology and taxonomy.

THESIS 2. An ontology is not reducible to pure cognitive analysis (in philosophical terms, it is not an epistemology or a theory of knowledge). Ontology represents the 'objective' side (= on the side of the object), and the theory of knowledge the subjective side (= on the side of the knowing subject) of reality. The two sides are obviously interdependent, but this is not to imply that they are the same (exactly like the front and rear of a coin). In order to conduct ontological analysis, it is necessary to 'neutralize', so to speak, the cognitive dimension, that is, to reduce it to the default state. I assume that the default state is the descriptive one, where the dimensions of attention, of interest, etc., are as neutral as possible (= 'natural' attitude). It is of course possible to modify the

default state and construct ontologies of the other cognitive states as well, but this involves modifications of the central structure.

THESIS 3. There is nothing to prevent the existence of several ontologies, in the plural. In this case too, ontological study is useful because, at the very least, it renders the top categories explicit and therefore enables verification of whether there are reasonable translation strategies and of which categorization can serve best to achieve certain objectives. This thesis is a subtle argument and requires some further comment. Let us assume that the general space of the ontology is a virtual space which is actualized in its concrete instances. The situation can be exemplified by citing the case of natural languages. The intended reference is obvious here and implies that the universal ontology is like the universal language. The least that one can say is that neither of them exists. Just as there is no universal language, so there is no universal ontology. What exists instead are individual natural languages, each of which is general; that is, it is able to say everything that it wishes to say. By the same token, there may not be a single universal ontology, but there may exist several general ontologies. Also in this case (which is the one that I regard as closest to the truth), structural studies of great interest are possible. For the sake of clarity, I shall continue with the example of languages. However different various languages may be, they display a number of highly significant uniformities. If we divide the structure of a language between lexicon (= content) and grammar (= structure), we note that there are dimensions which, with by and large limited variability, are systematically treated by the lexicon or by the grammar. For example, some languages have singular and plural, others add a dual form, others a trial form, others can grammatically indicate the many and the few. But that is all. No language has grammatical devices with which to denote the countable or 435. In these cases they resort to the lexicon. What I wish to say is that, for all their differences, languages display certain structural uniformities. The same should apply to ontologies. However different the various general ontologies may be, some structural uniformities among the top categories should nevertheless exist [Talmy 1988]. If this argument holds, two prime focuses of attention ensue from it:

- (a) the construction of a general ontology to be used as a test framework;
- (b) the theoretical and experimental analysis of the process of multiple categorization.

The first point is obvious. Without a general ontology we have no context of reference. The second is much more delicate. If the categorization is intended to be a passage from a substratum space to a target space (possibly discrete), the problem becomes that of translating objects categorized according to the categories of one particular target space into objects categorized according to the categories of another target space. One of the reasons for the difficulty of the problem is that we often find that we know only target spaces and not the substratum space whence they originate, of which we must hypothesise the characteristics – as far as we can – from the target space or spaces [Petitot 1985].

THESIS 4. Reality is organized into diverse levels and there are highly sophisticated interdependences among these levels and within them. For instance, an adequately sophisticated ontology must be able to perform the following two tasks at the very least:

(a) distinguish among such diverse objects as the things or inanimate concrete objects of the material world, organisms or animate concrete objects, the psychological objects of the mind, the social and institutional objects of social life, the abstract, ideal and fictitious objects of the sciences and arts.

(b) connect the various levels of reality by means of opportune forms of dependence. A promise as a social act with juridical force requires the existence of people with juridical capacity, and that people do not fluctuate in some hyperspace but possess living bodies, which are in turn the result of particular supraformations of material entities [Hartmann 1966].

3. Levels of reality

The universal glue of any whatever articulated ontology is given by the network of dependencies that it creates and governs. We may distinguish at least five ontological levels: of the inanimate physical world, of the animate physical world, of the psychological world, of the social world and of the ideal world.

Specific forms of categorial and existential dependence exist among these levels. For example: a psychological object 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. Hypothesising forms of existential dependence does not signify resorting to more or less overt hypotheses of reductionism. The various ontological levels may be existentially constrained without this implying that they are categorially constrained. The categories or properties which enable us to describe the world of psychological states are different from the categories that enable us to describe the world of animate objects and that of inanimate objects.

In more analytic terms, we may hypothesise that (a) new categories may intervene in every passage from one level to another, and (b) the lower-level categories that reappear at a higher level undergo transformations, they are 'supraformed'.

The supraformation relationship is that which inheres, for example, in the passage from the inanimate material world to the animate material world. The categories of the former maintain their validity also for the second, but in a different form which reflects the specificity of the new level.

Differing from the supraformation relationship is the supraconstruction relationship, which holds, for example, in the passage from the animate material world and the psychological world. In this case, the higher level requires the lower one only as its *external basis of existential support*, but not as material to be supraformed (in cases like this one, the substratum of the higher level is not the material of the lower level).

There are several kinds of dependence relation among the various levels. For example, every higher level not only requires corresponding lower levels but also that these must have specific structural characteristics (of complexity, for instance).

In general, we may distinguish the general categories, which hold for all the ontological levels, from the regional categories, which apply only to certain levels. The general categories comprise, for example, object, event, substratum, substance, form, relation, determination, dependence, structure, part, whole, unity, multiplicity, dimension, continuum, discrete, internal, external, identity, diversity, possibility, actuality, necessity, change.

Examples of regional categories are the following. For the material world: space, time, cause, situation, reciprocal action, dynamic structure, dynamic equilibrium, becoming. For the animate world: organic structure, adaptation, end-directedness, material exchange, self-regulation, life of the species, degeneration of the species. For the psychological world: act, content, consciousness, unconsciousness, pleasure, displeasure. For the social world: social system, family, community, conflict, class, institution, integration. For the ideal world: the categories of the activities and products of knowledge, art and faith.

These lists are obviously provisional and their purpose is purely exemplificatory. A thoroughgoing ontology would have to refine them by making appropriate changes and studying the connections among the various categories.

This proposal for a stratified construction of ontology can be summed up in the following theses:

1. The general categories encompass all the various ontological layers. However, they display features at every level which are specific to that particular level (because they interact with the complex of categories at that level).

2. The categories of the lower ontological layers are the foundation for the higher ones, but they are independent from the higher categories.

3. The categories of the lower layers are stronger than the categories of the higher layers, but they have lesser structural power.

4. In the case of supraformation, the whole categorized by the lower categories helps to constitute the substance of the higher level.

5. In the case of supraconstruction, the whole categorized by the lower categories acts as the existential bearer of the higher level.

4. The trap of simplicity

An epistemological thesis which finds wide acceptance is that of theoretical simplicity. It is often said that, on the assumption that a procedure exists for calculation the complexity of a theory, of two theories T1 and T2 relative to the same domain the simpler one is preferable. However obvious this thesis may be, it is potentially highly dangerous for any enterprise of an ontological nature.

The reason is that the central problem of ontology is not so much the categorization of any domain whatever as the categorization of the maximal domain. What is needed, therefore, is not the simplest theory for any domain whatever, but *the simplest theory for the maximal domain*. Accordingly, it is almost certain that the simplest theories for local domains are also those which most resist expansion to embrace broader ones. From an ontological point of view, the most useful theory is the one that can be expanded in the most economical manner possible.

We can reformulate the thesis by stating that it is advisable to spend more on constructing local theories because in this way less need be spent later when integrating and coordinating them.

If we take any two basis of data, there is no guarantee that their data will be interchangeable. The ontology (once constructed) must constitute the context into which the two data banks are merged, thereby modifying the data in order to render them interchangeable. The translation will be simpler, the more adequate is the structure of the initial basis of data.

5. A word of caution

In constructing an ontology it is important not to confuse ontological problems with those involved in their formal translation. In other words, care must be taken to distinguish between the ontological tree and the logical tree that should be its rigorous translation. I see at least three reasons for keeping the two trees distinct. First, the categories that constitute the two trees are not the same, and in passing from one tree to the other changes may occur of which one should be aware, also because there is nothing to guarantee the neutrality of the translation. For example, from the ontological point of view, 'property' and 'relation' are very different entities. From a logical point of view, however, it is obvious that the former is a monadic relation. This type of translation, therefore, may be anything but innocuous.

Another reason for keeping the level of ontological categories distinct from that of the logical categories is that there may be different logical translations of the same ontological structure. These different translations may prove to be entirely compatible; but they may equally be incompatible and usable in different ways according to the particular purposes of the moment.

Thirdly, at the actual state of research, apart from the fact that there is no reason for believing in a universal ontology, there is nothing to guarantee that one single logic applies to the entire ontology. In other words, there is nothing to rule out that the different sections of the ontology may

not be more efficaciously formalizable using different logics. In short, the ontology and the logic (or at least the formalism) which should give it formal rigour lie at different levels which should not be confused. Both must be addressed, but each *iuxta propria principia*.

6. The perspective of the whole

On the basis of the foregoing discussion it is now possible to outline a hypothesis of ontology.

I assume the category of whole as the top level. Synonyms frequently used for this category are: entity, thing and object. I prefer whole because it highlights from the start a crucial aspect which is not stressed with equal emphasis by other proposals: being a whole means having boundaries. Boundaries which may be more or less rigid and of different kinds. The type of boundary that may be relevant on any particular occasion depends on the level of the whole. Although inanimate, animate, psychological, institutional, abstract, imaginary, etc. wholes may have different boundaries, they always have boundaries.

By virtue of possessing boundaries, a whole is something on the basis of which there is an interior and an exterior. Put in different but not alternative terms, we may also say that a whole is something which displays some form of independence with respect to an environment. Observing that, when analysed at a sufficient level of detail, every whole vanishes into a continuum, or according to which every whole depends on something else, does not raise major difficulties. The fact that the boundaries of the whole are not absolute does not imply, in fact, that these are purely apparent boundaries. Wholes and their boundaries are realities which effectively operate at the appropriate level of granularity.

In dealing with wholes composed of other wholes, the problem arises of calibrating the 'weight' of the more general whole with respect to the 'weight' of the boundaries of its component wholes. For these cases I propose the following hypothesis: for wholes constituted by several wholes: the boundaries of the wholes of the ontologically lower layers are *stronger* (= more stable) than the boundaries of the wholes of the higher ontological layers. Both the sheep and the flock of sheep are wholes, but the boundaries of the individual sheep are stronger/more stable than those of the flock.

My analysis of the whole now proceeds by drawing on two groups of theories: those that give information about the whole as such, and those which give information about the various ways in which the whole can be considered.

7. The whole as such

Analysis of the whole as such examines level, location and amount.

By 'level' I mean the ontological layer of the whole and therefore whether it is an inanimate material object, an animate object, a psychological object, a social object or an ideal object. This classification activates the appropriate forms of existential dependence and the sets of specific categories which hold at that particular layer of reality.

By 'location' I mean the position of the whole in one of the semantic fields of space, time, matter and substratum (or existential bearer). These four semantic fields are appropriately activated by the level. For example, in the case of wholes of inanimate material type, one finds that matter and substratum coincide; whereas certain ideal wholes do not have a temporal location, etc.

'Amount' tells us that we are dealing with only one whole, with several wholes, or with material denoted by mass terms. I assume that the default case is the one involving a single whole. If instead several wholes are given, then new categories are activated, like those of relation and force. By way of a further example: significant applications of the category 'relation' are those which

concern the relations among wholes in terms of their location. In the case of two wholes, for example, relation < pattern < spatial < schematic < behind/in front; above/below; right/left; diagonal. Etc.

The appropriate categories are activated in the case of matter denoted by mass terms.

8. Modes of considering the whole

The whole has parts – which I distinguish into separable and non-separable – and a history.

Let us first inspect the distinction between separable and non-separable parts. However difficult it may be to specify this distinction, a preliminary definition of the difference may be forthcoming from examination of the case of inanimate material objects. In this situation, we may call separable parts those which can be removed from the whole without anything else taking their place. We may therefore say that a leg is a separable part of a table. The removal of separable parts may have a destructive effect on the whole. Non-separable parts are instead those which can be recognized and distinguished but which cannot be removed [Husserl 1970, 3rd Investigation]. When they are ‘removed’ from the whole, a part of the same kind usually takes the place of the part that has been removed. In effect, what is really inseparable is not so much the individual part *qua* part as its genus. If I remove the colour ‘red’ from a table, it becomes of another colour or it assumes another colour, but it is still in some way coloured. For categories like colour, weight, shape, size, consistency, etc., material objects are structured in such a way that they may display different instances of these categories in the course of their histories, but they nevertheless always have them. A table may be of different shape, colour, weight, consistency, etc., but it will always have a certain specific shape, colour, weight, consistency, etc.

With the appropriate variations, the difference between separable and non-separable parts holds for every type of whole.

The different ways in which separable parts are given enable us to provide a classification of the types of whole. We may accordingly speak of systems when the separable parts are given simultaneously and the boundaries of the parts are less strong than the boundaries of the whole. We may instead speak of aggregates when the separable parts are given simultaneously but the boundaries of the parts are stronger than the boundaries of the whole. We may then speak of events when the parts are given in succession and the boundaries of the parts are weaker than the boundary of the whole; reserving the term process for the case in which the parts are given in succession and the boundaries of the parts are stronger than the boundary of the whole.

As for the history of the whole, I distinguish between the history of the whole as such and the history of the location of the whole, according to the modification and/or substitution of its parts. The histories relative to the location of the whole and to the modification or substitution of its parts enable us to establish various criteria of identity (material, functional, etc.).

9. Conclusions

I have tried to show that constructing a general ontology requires the attention to be focused on – besides the problem of the theoretical and experimental analysis of multiple categorization (§ 2) – the theory of dependence. My ontological proposal has been an attempt to articulate at least two of the many and diversified forms of dependence among wholes (as dependencies among ontological layers and among levels of one layer) and internally to one whole (as dependencies between the whole and its various kinds of parts, and between the whole and its history and that of its parts) [Ingarden 1964].

References

- Burkhardt, Hans and Smith, Barry, eds. (1991). *Handbook of Metaphysics and Ontology*. Munich: Philosophia.
- Guarino, Nicola and Poli, Roberto, eds. (1995). *Formal Ontology in Conceptual Analysis and Knowledge Representation*. Monographic issue of *International Journal of Human-Computer Studies*, 43, 1995.
- Hartmann, Nicolai. (1966). *Zur Grundlegung der Ontologie*. 3rd edition. Berlin: W. de Gruyter.
- Husserl, Edmund. (1970). *Logical Investigations*. London: Routledge and Kegan Paul (2 vols.).
- Ingarden, Roman. (1964). *Der Streit um die Existenz der Welt*. Vol. 1: *Existentialontologie*. Tübingen: Niemeyer.
- Petitot, Jean. (1985). *Morphogenèse du sens*. Paris: Presses universitaires de France.
- Poli, Roberto. (1992). *Ontologia formale*. Genova: Marietti.
- Poli, Roberto and Simons, Peter, eds. (1996). *Formal Ontology*. Dordrecht: Kluwer.
- Talmy, Len. (1988). *The Relation of Grammar to Cognition*. *Topics in Cognitive Linguistics*, B. Ruzska-Ostyn ed., 165-207. Amsterdam-Philadelphia: John Benjamins Publishing Co.
- Sowa, John. (1995). *Top-Level Ontological Categories*. *International Journal of Human-Computer Studies*, 43, 1995. Monographic issue on *Formal Ontology in Conceptual Analysis and Knowledge Representation*, ed. by Nicola Guarino and Roberto Poli.