Knowledge Management from the Ontological Standpoint[1]

Pierre Grenon

Institute for Formal Ontology and Medical Information Science
University of Leipzig
Haertelstrasse 16-18
D-04107 Leipzig
pgrenon@ifomis.uni-leipzig.de

Abstract: What is an adequate philosophy for knowledge management for an ontologist working in information science? I will contend that the relevant part of knowledge management, namely the theory of knowledge representation, lacks proper foundations. In particular, I will criticize the traditional philosophical approach to knowledge representation, which I call pragmatist conceptualism. I argue on the one hand that its emphasis on pragmatism falls short of providing a reflective and principled methodology, and that on the other hand conceptualism is doomed by its neglect of reality. This paper outlines and defends a clear standpoint of philosophical realism together with a commensurate methodology for the task of the knowledge engineer.

1. Introduction

Knowledge management is a broad field of researches and applications drawing on expertise from a large variety of domains. The driving force of knowledge management could be presented as an attempt to produce efficient and re-usable tools for the understanding and manipulation of human and machine-processable knowledge. Knowledge management could thus be described as aiming to provide a framework for anticipating the unknown. The objects of knowledge management are belief systems, models, texts, theories, and like resources. There is in addition one pervasive and centrally important type of manipulation of knowledge in the field of machine processing, namely that which is involved in the construction of databases. I will take the standpoint of the builder and maintainer of knowledge in a database, called in this context a knowledge base.

The task of the knowledge engineer is that of putting knowledge into a computer-processable form. Typically this is done by imposing upon the data a formalism that enables it to be stored and manipulated. This forms the core of the knowledge representation task. Storing and handling information are tasks that belong to the data

[1] This paper is based upon work supported by the Alexander von Humboldt Foundation under the auspices of its Wolfgang Paul Programme. I am indebted to Ingvar Johansson, Pierluigi Miraglia, Michael Pool, and Barry Smith for their helpful comments.

management and information retrieval part of knowledge management. The resultant formal representation is often used to provide support for applications such as reasoning or natural language processing, which are in turn sometimes considered to fall within the scope of the knowledge engineer's activity.

There is a variety of techniques and formalisms that may be used by the knowledge engineer. I will take as paradigmatic[2] the representation of knowledge through a logical formalism such as that of first-order predicate calculus. This has the advantage of allowing in principle the explicit representation of the objects in the relevant domain of discourse. Against this background, I would like to discuss the use of philosophy for knowledge management and in particular for the methodology of knowledge representation.

There is not one philosophy, but rather many philosophies. I make the fundamental assumption that some philosophies are more useful than others for a given purpose. In particular, this is true of the range of philosophical positions which may underlie an approach to the practice of knowledge representation and by extension to the field of knowledge management. I wish to question what can be regarded as the traditional philosophical approach to knowledge representation, namely a joint venture of conceptualism and pragmatism, and extract from a critical analysis thereof some methodological maxims which might prove useful to knowledge engineers. My ultimate goal is to find and expose the best philosophical position for the purpose of knowledge representation, and I will defend to this end a rich realist approach. I believe that any result will extend to knowledge engineering and knowledge management in general. I will first introduce the practice of knowledge engineering and the use of philosophy therein, before giving details concerning the most relevant philosophical field, namely ontology. I will give a critical exposition of what I regard as the mainstream philosophy of knowledge representation, namely pragmatic conceptualism. Next, I shall attempt to warn against certain fatal tendencies of the conceptualist approach. Finally, I will present a set of philosophical positions (namely realism, perspectivalism, adequatism, and fallibilism) which in my opinion form an adequate basis for the methodology of knowledge representation.

2. Ontology and Knowledge Representation

Knowledge engineers structure bodies of information. Their activity is shaped *inter alia* by philosophical assumptions underlying their adopted methodology, and also by presuppositions inherent in the specific domain or implicit in the structure of the framework in which the formalization of information is conducted.

The work of representation begins after the tasks of knowledge acquisition or information gathering have been completed. It can be broken down into three overlapping main tasks: i) feeding the database, ii) improving the existing framework, and iii) formalizing the

[2] See [BSE90] for a somewhat dated but remarkable introduction to the field of knowledge representation and its techniques.

knowledge of an expert in whatever is the pertinent subject-matter. The knowledge engineer builds theories, typically proceeding by generalization and abstraction, and sometimes even makes use of metaphysical (high-level philosophical) and ontological insights.

Philosophical ontology is a branch of philosophy which concerns itself with the question of what there is, i.e., what are the entities existing in the world and what are the categories under which they fall? The product of an ontological investigation is typically a taxonomy of categories. In addition, ontology will provide an account of the relations between entities and of the structure of the world at a very high level of generality.
Husserl in his [Hu13] has made popular the distinction between at least two kinds of ontological inquiries. On the one hand is formal ontology, which conducts analysis and produces theories at the highest and most domain-neutral level possible, the level of *forms*. On the other hand, is material or regional ontology, which is the ontology of some specific domain or material region.

The term 'ontology' in information science has very many uses. An ontology has been defined, for instance, as:

1. a set of terms (classes, categories, concepts, words)[3],
2. an axiomatic theory or a set of propositions (see [MeForth.]),
3. the 'content' (conceived in a rather loose sense) of a knowledge base in general (or of some specific knowledge base such as that of Cyc).

There is one predominant view which is consistent with most uses of the term, namely the view that ontological engineering is a form of modelling. Indeed some even speak of ontology as a matter of conceptual modelling or of conceptual representation. The knowledge engineer becomes an ontological engineer[4], I would like to propose, when performing a philosophical analysis of the content and a shaping of the infrastructure of the knowledge representation system in the light of metaphysical/ontological theories.

3. Trends in Knowledge Representation

I hold a position which I will call realist representationalism. According to this position, the basis for knowledge representation should be, not representations of reality, but rather reality itself. This is in contrast to the mainstream approach to knowledge representation which understands the task of the knowledge engineer as consisting *only* in that that of representing others' (the domain experts') representations (so that reality falls out of the picture almost entirely). Knowledge is conceived as a matter of mere conceptualizations, and there is no care given to whether it really represents a reality that is independent thereof. Thus, the task of knowledge representation becomes that of providing a formal

[3] Such sets can be more or less structured and can be anything from taxonomies (see for instance [SmForth.]) to dictionaries, vocabularies, terminological systems or even thesauri.
[4] The distinction between knowledge engineering and ontological engineering essentially amounts to a division of labour. This distinction was made already by Russell and Norvig (in their [RN95]) for whom it corresponds to the opposition between the domain-specific and the domain-neutral.

framework for the representation of a conceptualization. In my opinion, knowledge representation should take one step closer to earth and assume the project of trying to depict reality itself. My position is thus a form of realism as contrasted with that of conceptualism.

Epistemological conceptualism can be defined as a view according to which there is no knowledge of reality but only of our own concepts. There are concepts in minds, and knowledge can only be knowledge of such concepts. Reality is always outside, beyond, transcendant to what we can know. Conceptualism with regard to universals says that general terms only refer to universals as concepts and not to universals in the world. Conceptualism gives primacy to the representation of concepts (or ideas) in its understanding of knowledge. It tends to give a major role to epistemology; to the detriment of both ontology and the reality that ontologists address. The mainstream approach to knowledge representation might well be called pragmatist conceptualism. According to this view, knowledge representation deals only with conceptualizations; the presumed reality is for other people (scientists, for example) to deal with. Therefore, the knowledge engineer is free to bring up any kind of objects and whatever theoretical constructs are required to the extent that such summoning logically fulfil the practical purpose of the representation. This is emblematically endorsed by Genesereth and Nilsson in [GN87]. The authors sum up their position as follows:

> no attention has been paid to the question whether the objects in one's conceptualization of the world really exist. […] Conceptualizations are our inventions, and their justification is based solely on their utility. (p. 13)

My position is that although it is perfectly proper for us as knowledge engineers that we should be looking for a useful account, the best way to realize this goal is precisely by looking for an account that is adequate to reality. This is, above all, likely to result in an improvement in stability of the resulting framework.

From the realist point of view, information science puts far too heavy a load upon the term 'concept'. A 'concept' is taken to be:

1. an idea or a mental representation of objects in reality;
2. a general idea under which there falls a multiplicity of things (I call such things conceptual universals);[5]
3. a Platonic[6] idea existing as a perfect prototype of things in the world, but itself exterior to the world;
4. a class;
5. a word;
6. the meaning of a word.

[5] Reflected in a language by the use of a (typically monadic) predicate; they allow great flexibility in representation, but it should be noted that only some correspond to any counterparts in reality. There is no universal in the world, for example, corresponding to our general concept of a unicorn.

[6] Here this term can be taken as synonym of abstract and perfect model.

As a perusal of the literature of information science will quickly reveal, these notions are often run together.[7] Many knowledge engineers do not notice the ambiguity in the use of the term 'concept', which in the literature of computer science is associated rather with the use of the term 'type' whenever some generality occurs. The pervasive use of 'concept' and the running together of its various meanings poses a thorny problem for anyone working within knowledge representation concerned with clarity of thought and clarity of expression. The distinction between things (or real entities) and the corresponding concepts is acknowledged by the knowledge engineer; in his actual work, however, the things are neglected to the benefit of their conceptual proxies. No wonder, then, that generality is given a cursory (set-theoretic, Platonic, linguistic, etc.) reading. In contradistinction, realism maintains in Aristotelian spirit that in addition to the general terms in language and the conceptual universals in our minds, there are also universals in re, i.e., universals really existing in the world (existing, on Aristotle's view, in their instances). These universals are not concepts; they are invariants in the things themselves.

I contend that pragmatist conceptualism lacks a principle of cohesion. It seems to be but a claim of freedom against any constraints deriving from reality itself other than that of (putative) utility. It provides us with no principle assuring us that a bottom-up (from domain-specific to domain-neutral) and a top-down (from domain-neutral to domain-specific) approaches will meet in any coherent way. There is also too little attention paid to a principled resolution of the problem of unifying the many different and often mutually incompatible idiomatic frameworks and representations developed independently by different groups.[8] For the conceptualist, if there is a multiplicity of 'conceptualizations' all should be accounted for. The question then only becomes one of how to relate (or 'fuse') them; and such an emphasis has the consequence that the question of adequacy to reality of the conceptual schemas tends to be neglected. The fashionable recourse to namespaces, using a syntactic device marking the origin of a 'concept', helps as a bookkeeping device but it eliminates only some of the confusion. It does not provide the needed ticket for interoperability. In effect, we end up having concurrent systems allegedly representing the same reality. What, if we wish to fuse these competing systems together, can serve as tertium quid? What, when concepts differ, is to serve as our guide in resolving the differences?

4. Tendencies in Conceptualist Knowledge Representation

There is in my opinion a door opened by pragmatist conceptualism to various tendencies toward confusion of thought, of which I will discuss three types, which I will call linguistic, subjectivist, and algebraist. These tendencies lead to the risk of succumbing to fallacies of linguistic and theoretical imagination.

The linguistic tendency rests on the conviction that natural language is already the best medium for representing reality (it is after all the most commonly used) and that any

[7] A puzzling example is provided by nothing less than an international standard such as ISO 1087-1:2000.
[8] This is the well-known problem of the Tower of Babel.

knowledge representation system ought to reach perfection by imitation of natural language. But we must question how much natural language should dictate an ontological inquiry. The risk is precisely a radical claim according to which an ontology would have to stand in a one-to-one correspondence (or as close thereto as possible) with the elements of natural language (one word, one 'concept'), and thus with an overwhelmingly rich system of concepts. At best, in my opinion, natural language can serve as a first clue to the ontologist, but it should certainly not be a criterion of correctness of the end-result of his labours. Indeed, if everything in natural language is in order as it stands, then there seems to be no need for ontological structuring or conceptual modelling.

A second tendency could be named subjectivism. According to this view, the world is the product of a subject's conceptualization. There are possibly as many conceptualizations as there are conceptualising agents. Thus, possibly as many 'ontologies', since for the conceptualist an ontology is but a set of (possibly consensual) conceptualizations. This raises the obvious problem of radical and permanent interpretation. In other words, it takes the Tower of Babel as a premise. This view is akin to relativism and a position defended in philosophy as perspectivism according to which no one possible perspective (or conceptualization) has more value than another. Such an assumption is already active in pragmatist conceptualism.

The last tendency I shall warn against can be called algebraism. There are actually two forms of algebraism depending on the objects considered: on the one hand, concepts, general terms in a language, or even intensional objects (which relate to issues of content), and on the other hand, theories (which relate to issues of structure). The common assumption, no matter the objects, rests on an ideology according to which a knowledge representation system ought to be maximally complete and contain all logical (algebraically, sometimes set-theoretically, constructed) *possibilia* (possible objects of manipulation). There is thus no limit to construction and all variants fall within the scope of the resultant knowledge representation system. That a total conceptualization can be viewed as a system of concepts with both an initial set and an articulating theory provides the root for the distinction between the two forms of algebraism. My brief against this tendency comes from a general scepticism in regard to the arbitrary production of fictions for the sake of systematic neatness, and the further habit of producing whole theories out of fictions. This leads to an explosion of the domain and to a huge problem of relevance and choice among theoretical possibilities and variants. Most of all, it does not account for the natural joints of reality, which (in a domain like medicine, for example) fall far short of algebraic neatness.

5. Toward an Adequate approach to Knowledge Representation

The three aforementioned tendencies of pragmatist conceptualism have this in common: that they do not take reality sufficiently seriously and that they do not assume the consequences of the fact that not all that is conceivable is thereby real or even possible. I argue that realist representationalism, i.e., an approach to knowledge representation based on reality, should be the basic methodological approach to the task of knowledge

representation, and more generally to the whole business of knowledge management. However, it needs to be complemented by three further foundational elements, namely perspectivalism,[9] adequatism, and fallibilism.[10]

Realism is first and foremost a claim about the existence of the world and its constituents. It is also a claim that the world and the entities it contains exist independently of our (linguistic, conceptual, etc.) representations thereof. Although in a sense less rich than conceptualism, realism is actually more focused and restricted to the most valuable target of our endeavours, namely reality itself. I endorse scientific realism which asserts that knowledge of reality can be obtained through scientific inquiry. Three maxims can be usefully retained at this stage: i) the primary source of knowledge is reality, ii) the domain of inquiry is a part of reality, iii) the expert knows the reality at hand (that is why we call him an expert).

Realist perspectivalism maintains that there may be equally legitimate realist perspectives on reality. It is important to bear in mind in this connection that this does not amount to the thesis that any view of reality is legitimate. To establish which views are legitimate we must weigh them against each other and against their ability to survive critical tests when confronted with reality itself, for example in scientific experiments. Those concepts and conceptualizations which survive are then transparent to reality. More generally, we are concerned with those views that are veridical under a given perspective, in relation to a particular domain, and at a given level of granularity (microscopic, mesoscopic level of everyday objects, geographic, macroscopic and even cosmic levels).

Adequatism is the opposite of reductionism. The adequatist affirms that there are many views of reality all of which are transparent. Each legitimate view has to be regarded as genuinely relating to reality, and the depiction of reality under a given perspective may be veridical even if some reductionist explanation proposes to do away with them.

To embrace scientific realism is to endorse the view that science can culminate in genuine knowledge of the world. Now, it is a fact that the sciences evolve and progress, and so does our mundane knowledge. It must be therefore that our theories and our understanding of the world can be subjected to trial, progress and revision.

6. Conclusion

My position can be summarized as follows: a philosophical and ontological analysis has to be performed to provide a sound basis to 'conceptual modelling'. More precisely, knowledge representation systems should be conceived as representations of reality and not as systems of representations of mere concepts or models. A complete system should be able to accommodate and articulate possibly many legitimate views of reality. Which

[9] I will call this realist perspectivalism in order to defuse the traditionally radical relativist reading of perspectivalism (or perspectivism).

[10] This is a combination which is already put forward in [SmiForth.] although realist perspectivalism is only implicit therein. It is the combination which forms the methodological background of the work of IFOMIS.

alternative theories or perspectives on reality are useful for the purposes of the knowledge engineer has to be established on the basis of a realist and fallibilist methodology, and this task, which comes close to the tasks of science, may be the most difficult and challenging to accomplish.

Bibliography

[BSE90] Bibel, W.; Schneeberger, J.; Elver, E.: Representation of Knowledge. In *Knowledge Engineering: Volume I, Fundamentals*. McGraw-Hill, New York, 1990.

[GN87] Genesereth, M.; Nilsson, L.: *Logical Foundation of Artificial Intelligence*. Los Altos, California Morgan Kaufmann, 1987.

[Hu13] Husserl, E.: *Ideen au einer reinen Phaenomenologie und phaenomenologischen Philosophie*, 1913. English translation by Boyce Gibson, W. R.: Ideas: General Introduction to Pure Phenomenology. Collier-Macmillan, London, 1931.

[MeForth.] Menzel, C.: Ontology Theory. In H. Stuckenschmidt (ed.), *Ontologies and Semantic Interoperability*, CEUR Workshop Proceedings, forthcoming.

[RN95] Russell, S.; Norvig, P.: *Artificial Intelligence: A Modern Approach*. Prenctice Hall, 1995.

[SmForth.] Smith, B.: Ontology and Information Science. In *Stanford Encyclopedia of Philosophy*, forthcoming.