The Differentia Principle as a Cornerstone of Ontology

Prof. Christophe ROCHE
Université de Savoie - Campus Scientifique
73 376 Le Bourget du Lac - cedex - France
tel : +33 (0) 4 79 75 87 79 - fax : +33 (0) 4 79 75 88 88
roche@univ-savoie.fr - http://ontology.univ-savoie.fr

Abstract

The Information Society relies more and more on the co-operation and collaboration of multi-disciplinary people who need to communicate and share information. Communication and knowledge sharing are the new economic stakes. But everyone speaks his own language, with his own terms and meanings. The information society is a Tower of Babel which has to evolve to the knowledge society.
Since the beginning of the nineties, ontologies seemed being as one of the most suitable solutions faced with this problem and became a very popular research topic in knowledge representation.
Nevertheless several problems remain which claim for clarification. As a matter of fact, there is no agreement about what an ontology is; and the numerous systems as well as not very clear epistemological principles are barriers to the real use of ontologies. The ontology problem requires a multidisciplinary approach based on sound epistemological, logical and linguistic principles.
This article presents the "differentia" principle, as it appears in Aritotle's works and Porphyry's Isagoge, as a cornerstone for building ontology. We shall see that the metaphysical point of view is a guarantee for ontology commitment, reuse and sharing.
This approach will be illustrated with the OK (for Ontological Knowledge) computational model and the associated ontology-oriented language LOK.

keywords: *ontology, differentia, Porphyry'Isagoge, knowledge representation, terminology, computational model.*

1 Introduction

The Information Society relies more and more on the co-operation and collaboration of multi-disciplinary people who need to communicate, share and exchange information [1], [2]. Communication and knowledge sharing are the new economic stakes.
The Information Society, virtual enterprise, e-business, etc. rely on communication between of interacting and heterogeneous actors : people, organisations and even software systems. The communication is difficult due to the fact that each of these actors speaks a different language. The information society is the new Tower of Babel which has to evolve to the knowledge society.

To address this problem, we need a common (communication) language that actors can read and understand. Using a single, normalised language like KQML (Knowledge Query Manipulation Language [3]) can reduce the gap of misunderstanding by using a same syntax. But, although such languages give some useful indications about the pragmatic content of the message (by using predefined performatives i.e. commitment actions), the semantic problem has still to be addressed. As a matter of fact, two actors can communicate only if they agree upon on the meaning of the terms they use.
Ontology, understood as an agreed vocabulary of common terms and meanings within a group of people, is a solution to the communication and knowledge sharing problem.

2 Ontology

Since the beginning of the nineties, ontologies have become a very popular research topic in computer science including knowledge representation and information management. Such a popularity is due to the fact that ontology provides a means of capturing a shared understanding of terms that can be used by humans and programs.

It is amusing to note that when the main goal of ontology is to normalise the meaning of terms, the term "ontology" itself is not clearly defined: “although *ontology* is currently a fashionable term, no agreement exists on the exact meaning of the term” and “seems to generate a lot of controversy in discussion about AI (artificial intelligence)” [4], [5]. In fact, ontology finds applicability in many domains of application in knowledge and software engineering, and each of them gives its own definition.

Let us see some examples from the general, and famous, Gruber’s definition: “An ontology is a specification of a conceptualisation” [4] to a more dedicated one: “The main purpose of an ontology is to enable communication between computer systems in a way that is independent of the individual system technologies, information architectures and application domain.” [6].

Nevertheless, although an ontology may take a variety of forms, it will include a vocabulary of terms and some specification of their meaning [6], [7], [8], [9].

The result is ontology has gained considerable popularity and several ontologies were defined, whether general or specific: Cyc, Mikrokosmos, Enterprise Ontology, TOVE, Sowa’s, etc. as well as ontology-oriented languages and software environments were realised: OIL, DAML, Protégé, KAON, Ontolingua…

But, several problems remain which claim for clarification. As a matter of fact, it is very difficult to reuse and share ontologies. For example, how can we combine the definitions coming from the two enterprise ontologies TOVE [10] and Enterprise Ontology [11] ?

The differences between the semantics of the systems, as well as not very clear epistemological principles, are barriers to the real use and re-usability of ontologies. And how can we trust in, and then use, an ontology which does not offer “any kind of guarantees” [12] ?

3 Objectives

Our main objective is to define the meaning of terms (words) which refer to conceptual knowledge (e.g. concepts used in a corporate knowledge as ‘turning’, ‘milling’, ‘stamping’… in a ‘mechanical machining’ ontology). It implies that we rely on the classical semiotic triangle [1] [“term” - <concept> - *example*] where the meaning of a term is the concept to which it refers to.

Building consensual and coherent ontologies seems a quite unreachable goal without clear and sound principles. It is the reason why we claim that such a problem requires a multidisciplinary approach:

- Linguistic, as we use words to communicate,
- Epistemological, since words refer to knowledge (here concepts) which represents their meaning,
- Logical, in order to guarantee some coherence.

In order to lay stress on the fact that the ontology problem is mainly a linguistic (we use words to communicate) and an epistemological problem (these words have a meaning), we shall set down our own definition of ontology:

> *« An ontology is a conceptualisation to which one or several vocabularies can be associated and which participates to the meaning of terms. Defined for a given objective, an ontology expresses a point of view shared by a community. An ontology is represented in a language whose theory (semantics) guarantees the properties of the ontology in terms of consensus, coherence, sharing and reuse. »*

Well, but what about the theory ?

4 Porphyry’s Isagoge

Words have a weight. Unlike some computer scientists [2], we think that we have to keep in mind that ontology belongs to metaphysics. Let

[1] from a linguistic point of view, one speaks of a combination of a “signifier” and a “signified” (Saussure’s structuralism).

[2] “The use of the term *ontology* is somewhat unfortunate since it has a definite meaning in the philosophical literature which has little to do with describing the content of information repositories”. KACTUS - Esprit Project 8145.

us begin from an etymological point of view. Ontology come from the Greek 'ontos', whose meaning is being, and from the Greek 'logos', whose meaning is both language and reason. So, when some people say that "Ontology is the branch of metaphysics that deals with the nature of being" and others "an Ontology is a systematic account of Existence[3]" it is not the same definition : Being and Existence are different.

Building ontologies is a difficult tasks, in particular if we want some guarantees like consensual definitions. It is necessary to follow clear guidelines based on sound epistemological principles. Let us give some of them :

- set and concept are two different notions. A set is not a concept, even if a concept can be understood as a set of individuals subsumed by the concept,
- valued attributes describe the state of things but do not define them,
- a concept is not a well formed formula: "*Being is. Being is not true or false*".

It is the reason why the Porphyry's Isagoge always remains a novel work for knowledge representation. The *quinque voces* (the five predicables): genus, differentia, species, proprium and accident can be considered as backbone principles for building ontology.

The fundamental idea is that concepts are organised according to the essence of things and not according to their state.

Differentia and accident allow to clearly identify essential knowledge when genus and species allow to identify and organize concepts.

5 The Differentia Approach

The approach by differentia[4] fulfils the criteria previously defined. In fact this approach is more epistemological and logical than linguistic. It focuses on the essence of objects rather than on what could opposite words[5].

5.1 Concept

If the concept is the meaning of term, its own meaning is clearly defined:

> "*a concept is defined by specific differentiation. It means that a concept is defined from a previously existing concept adding a new differentia which is then called the 'specific differentia' of the newly created concept*".

Such a recursive definition tightly links concepts and introduces an important new notion: the differentia. The "specific differentia" relationship between two concepts is more than the classical "is-a" relationship since it introduces constraints between the sibling and subsumed concepts. Let us notice that the meaning of a concept can also be defined by the set of its differentiae.

5.2 Differentia & Accident

Differentiae are the elementary units from which the meaning of a concept is built. This means that the differentiae have no meaning in themselves. Unlike an attribute a differentia cannot be removed from the definition of an object without changing its nature; nor be valued. For example, for the 'Turning' concept, 'piece rotation' is a differentia whereas 'rotation speed' is an attribute.

A differentia is an unit from which meanings are built and which divides concepts into two no connected sets. In fact, adding a differentia to an existing concept creates two new ones, the first one to which the differentia belongs to and the second one which will never be able to own that differentia. That differentia is called the "specific differentia" of the former new concept. The fact that a differentia cannot belong to a concept is itself a specific differentia which can also be named. This is the reason why differentiae are defined by couples of opposite differentiae, like 'metal

[3] For AI systems, what "exists" is what can be represented.

[4] we prefer to use "differentia" (differentiae in plural) rather than "difference" which is too vague in English when differentia is a very specific thing.

[5] unlike Saussure's Structuralism which opposes rather than to define. Nevertheless combining structuralism and differentia is fruitful: arbitrary nature of signs, meaning structure, finite number of differentiae as elementary units of meaning.

preservation' and 'no metal preservation' or 'material removal' and 'material deposit'. Thus, owning a differentia for a concept implies it will never contain the opposite differentia, nor the concepts it could subsume. Such a property will be very useful for building and using ontologies. It is a guarantee of the coherence of the ontology and forbid multiple inheritance.

The meaning of concepts is then structured into binary trees based on couples of opposite differentiae. The specific differentia of a concept is the opposite specific differentia of its sibling, and vice versa. In fact, the concept tree is less important than the sets of differentiae which define the concepts. As a matter of fact, the differentia is the elementary unit of meaning and the concept names are arbitrary. Furthermore, the same sets of differentiae can be represented by different binary trees as the order of differentiae in a set does not matter (a 'Human Being' is a 'Being' which is *mortal* and *reasonable* or a 'Human Being' is a 'Being' which is *reasonable* and *mortal*). So, we can conclude that: "*the meaning of a term is defined by a set of differentiae*".

Such a definition and the fact that differentiae are defined by couples of opposite differentiae imply a logical semantics of term meaning. This logical semantics is a guarantee of coherent ontologies and it is very useful for applications: information retrieval, acquisition...

Last but not least, the agreement problem is reduced to the sole problem of agreement on the terms that denotes differentiae. It is all the simpler since the application domain is technical and the differentiae are 'visual' and 'concrete' like 'tool rotation', 'no metal preservation', etc. and since few couples of opposite differentiae (let us say n couples) can define a lot of concept names (until to 2^n names).

5.3 Properties

The definition of concepts by specific differentiation could appear quite restricting, but in fact it offers several good properties which explain such a choice:

- everybody can agree with this definition: an 'Electro-plating' is an 'electro chemical' 'Machining';
- the differentiae are the elementary units of meanings. This implies that the agreement problem is reduced to the sole problem of agreement on the differentia names and not on the concept names ;
- no multiple hierarchy and therefore no problem of inheritance of different values;
- sound logical properties which are exploited during the building of ontology and which are also exploited in applications.

6 The OK Model

A computational model based on differentia theory, called OK for Ontological Knowledge, was defined as well as a dedicated language, called LOK, and a software environment, the OK Station©.

6.1 The LOK language

The LOK language (Language for Ontological Knowledge) is an ontology-oriented language composed of more than 150 instructions, with a 'à la Lisp' syntax. Those instructions are structured into two sets. The first set contains all the necessary instructions for definition and modification of 'term-meaning' couples that constitute the ontology. Let us take an example extracted from the following simple ontology about machining (see figure 1).

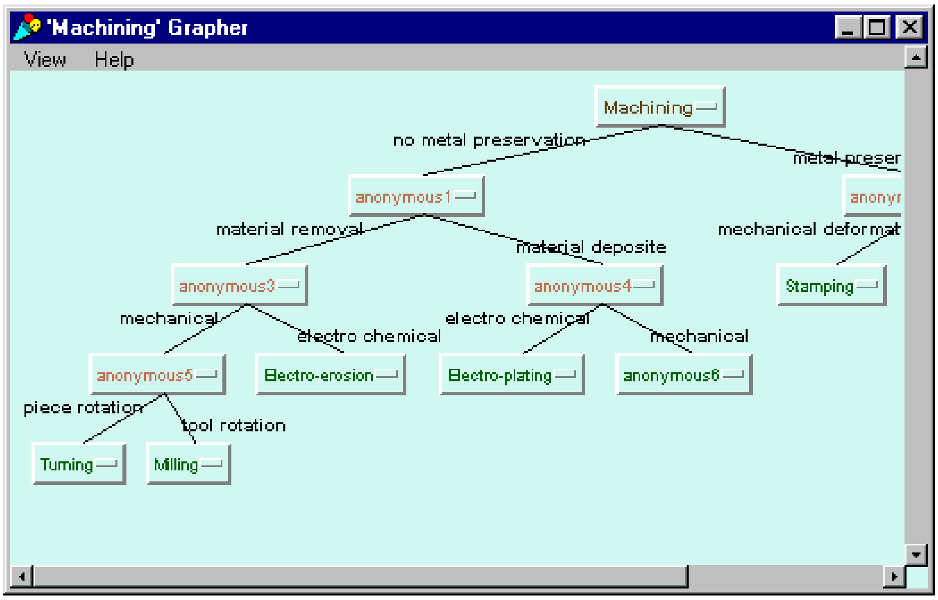


figure 1. a simple example of ontology

The building of this ontology begins by the definition of the differentiae using the ‘defineDifference’ LOK function:

(defineDifference
'metal preservation’),
(defineDifference
‘piece rotation’ ‘tool rotation’),

The last instruction creates two opposite differentiae whereas the ‘(defineDifference ‘metal preservation’)’ instruction creates only one differentia. In this latter case, the opposite differentia is automatically created with the same name and the prefix ‘a-‘ (‘a-metal preservation’).

As a concept is defined from a previously existing one, the first thing to do is to create a root concept. This is done with the ‘defineRootConcept’ instruction. For example the following LOK instruction :

(defineRootConcept
‘Mechanical Machining’)

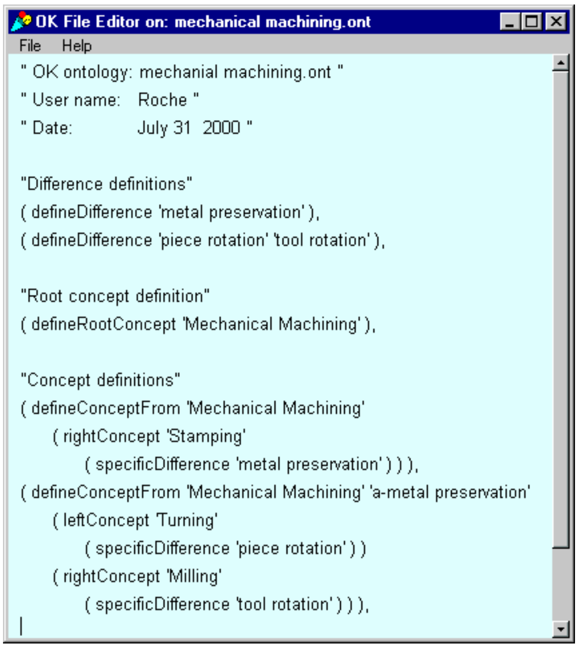
creates a new term, the string ‘Mechanical Machining’, without meaning because it is the root.

After that, new concepts can be defined. This consists in giving a new term and its meaning built from the meaning of an existing concept along with a specific differentia:

(defineConceptFrom 'Mechanical Machining'
(rightConcept 'Stamping'
(specificDifference
'metal preservation')))

creates the new term ‘Stamping’, whose meaning is the meaning of ‘Mechanical Machining’ plus the specific differentia ‘metal preservation’.

This simple mechanical machining ontology is defined by the following LOK file (figure 2):



```
OK File Editor on: mechanical machining.ont
File   Help
" OK ontology: mechanial machining.ont "
" User name:   Roche "
" Date:        July 31  2000 "

"Difference definitions"
( defineDifference 'metal preservation' ),
( defineDifference 'piece rotation' 'tool rotation' ),

"Root concept definition"
( defineRootConcept 'Mechanical Machining' ),

"Concept definitions"
( defineConceptFrom 'Mechanical Machining'
    ( rightConcept 'Stamping'
        ( specificDifference 'metal preservation' ) ) ),
( defineConceptFrom 'Mechanical Machining' 'a-metal preservation'
    ( leftConcept 'Turning'
        ( specificDifference 'piece rotation' ) )
    ( rightConcept 'Milling'
        ( specificDifference 'tool rotation' ) ) ),
```

figure 2. a LOK file

The second set of instructions are exploiting instructions mainly for queries. For examples:

(allConceptsWithDifference
'a-metal preservation')

returns the two terms ‘Milling’ and ‘Turning’ whereas : (listOfLeafConcepts) returns the terms ‘Turning’, ‘Milling’ and ‘Stamping’ as they are the only concepts from which new ones can be defined (leaf concepts). All these instructions are used by the different modules of the OK Station.

7 Conclusion

If ontologies have gained a large popularity in many domains of application as one of the most suitable solutions faced with the communication and knowledge sharing problems, several problems remain. As a matter of fact, an ontology will really be used only if everybody agrees on it and if some consistency is ensured. Using a same language is not sufficient, clear linguistic and epistemological principles are needed in order to reach a real ontology commitment. Ontology problem requires a multi-disciplinary approach.

We claim that the metaphysical approach of ontology must be kept in mind and that the Porphyry’s Isagoge always remains a novel work for knowledge representation. The *quinque voces* (the five predicables): genus, differentia, species, proprium and accident can be considered as backbone principles for building ontology. The fundamental idea is that concepts are organised according to the essence of things and not according to their state.

At last, we presented a model for ontological knowledge called OK (Ontological Knowledge) dedicated to conceptual knowledge, i.e. to the meaning of terms denoting abstract knowledge (concept). This model is based on the “specific differentia” theory which relies on sound principles taken into account linguistic and epistemological notions which may appear quite restricting. But such an approach where a concept is defined by specific differentiation provides many advantages: consensual definitions and logical properties which ensure consistency which are guarantees of ontology commitment, reuse and sharing.
The OK model has been implemented. The result is the OK Station©, a software environment dedicated to building, defining and exploiting ontologies. It is currently used to define ontologies in various areas.

References

[1] G.R. Olsen, M. Cutkosky, J.M. Tenenbaum, T.R. Gruber. “Collaborative Engineering based on Knowledge Sharing Agreements”, ASME Database Symposium, Minneapolis, 1994

[2] T.R. Gruber, J.M. Tenenbaum, J.C. Weber. “Toward a Knowledge Medium for Collaborative Product Development”, Artificial Intelligence in Design Conference, Pittsburgh, USA 1992

[3] Y. Labrou, T. Finin “A proposal for a new KQML Specification”, Internal Report TR CS-97-03, Computer Science and Electrical Engineering Department (CSEE), University of Maryland Baltimore County (UMBC), February 1997.

[4] T.R. Gruber, http://www-ksl.stanford.edu/kst/ what-is-an-ontology.html

[5] Nicola Guarino “Ontologies and Knowledge Bases. Towards a Terminological Clarification” Towards Very Large Knowledge Bases, IOS Press, Amsterdam, 1995

[6] http://www.ontology.org

[7] http://www.ontoknowledge.org/oil/

[8] M. Uschold, M. Gruninger. “Ontologies: Principles, Methods and Applications” Knowledge Engineering Review, vol ; 11, n°2, June 1996

[9] Mike Uschold and Martin King “Towards a Methodology for Building Ontologies” internal report AIAI-TR-183, July 1995

[10] M.S. Fox, The TOVE Project: Towards a Common Sense Model of the Enterprise, in Enterprise Integration, C. Petrie (Ed.), Cambridge MA: MIT Press, 1992

[11] M. Uschold, M. King, S. Moralee, Y. Zorgios, “The Enterprise Ontology” AIAI, The University of Edinburgh, 1997

[12] K. Mahesh “Mikrokosmos” http:// crl.nmsu.edu /Research/Projects/mikro/htmls/ontology-htmls /onto.index