Category Theory Praxis – The Semantic Interoperability Case
Charalampos Meletis, GR

Treating Semantic Interoperability Problems

This paper proposes the use of Category Theory, a branch of pure Mathematics, as an appropriate Framework for dealing with Semantic Interoperability issues.

The proposal is based on the author's Ph.D. Thesis, where as a mathematician and lover of abstract algebraic structures, he formulated a theory based on Categories, concerning Cognitive applications of Cybernetics, such as Analogical Concept Formation, Learning etc. These application cases have been treated considering the concept of Analogy as a mapping of semantic and context related characteristics between his Category-like structures which represent concepts of knowledge domains.

Thus, on the occasion of the present Conference, which coincides with the 30th Anniversary of my Ph.D., I'm taking the opportunity to put forward the above proposal of the innovative application of Category Theory as a research tool for treating Semantic Interoperability problems. A possible application to Knowledge Management research is also envisaged.

Prolegomena

Interoperability describes the ability and the tools needed for direct communication and synergy between different information and communications systems and organisational units based upon common standards, technologies and concepts.

Interoperability demands a shared understanding of information and an adjustment of data structure, which on the technical level means that heterogeneous devices (e.g. a mobile phone and a computer) can communicate to each other via a commonly agreed protocol (e.g. Bluetooth).

Semantic Interoperability (SEM-IOP)

Semantic Interoperability exists if in a data exchange between two systems, the data is interpreted in the same way by both systems ruling out misunderstandings and excluding misinterpretations which lead to semantic conflicts.

In order to ensure Semantic Interoperability, a standardised language is needed or a commonly agreed platform is required which enables two systems exchange information without any kind of human intervention.

The purpose of the communication between two systems is to carry out all the work needed in harmonised synergy as, for example, in the case of discovering, synthesising and delivering to a citizen a requested eGovernment service at national (NEGS) or at pan-European (PEGS IDABC) level.

To provide this sort of semantic consistency in Public Administrations applications at national and pan-European cross-boarder cross-sector (meta)level we need to develop and use at all levels a common language which amounts to the definition of a unified form of data representation and an appropriate semantics.

We must ensure that meta-information is described in such a way that it is interpreted consistently by all parties, as for example in the simple case of whether the <first name> field of an address record may contain several names or just one.

Among some proposals that have been made so far, including new ideas regarding SEM-IOP, we quote from (Rossiter et al 2006) who in his work 'A Natural Basis for Interoperability' argues:

''1. Need for Formal Natural Multi-level Type Systems.

Interoperability needs natural techniques to deal with levels of types. To handle (non-local) interoperability, formality (for reliability and predictability), naturality (for reality) and multi-level types (for types of types) are all required.

Categorical methods should replace classical models because models are local and interoperability is non-local. Categorical methods provide formal definitions of levels (as categories), mappings between levels (functors between categories) and comparison of one mapping between levels with another

(natural transformation between functors). Categorical techniques are also natural: an arrow

within a category is defined as unique up to natural isomorphism.'', and

''4 Discussion.

One of the purposes of developing a formalism for a problem area is to provide a rationale in which standards can be planned and discussed. It is perhaps only in the ideal world that standards are based entirely on a theoretical basis. Nevertheless some of the idiosyncrasies and inconsistencies of SQL have been attributed to not rigorously applying axiomatic set theory to the standard.

Category theory is a promising candidate as a formalism to assist in the preparation of an interoperability standard because of its pedigree as a workspace for relating different mathematics. The work here has shown that it can indeed perform this role with information systems and cover three critical areas of data structuring, constraints and manipulation (process) in an integrated manner. Recent advances in category theory are likely to improve its match with reality: 2-categories enable some of the strict criteria for composition and associativity to be relaxed to some extent.''

Furthermore, I would like to mention that despite that in the existing SEM-IOP initiatives, there is an agreement to use XML as the framework for data exchange, this does not enable IT systems to communicate each other and interoperate, in the same way as 'stringing together correct words of a language does not necessarily make a meaningful sentence'.

In view of the above considerations a question which arises to me is: are we on the right track (theoretically and methodologically) towards the ultimate goal of achieving a most globally accepted methodology and solution to the Semantic Interoperability issue?

Taking all these into account I would propose to look very carefully to Category Theory as an appropriate framework due to its inherent:

- data/content representation generality and clarity

- semantic virtues and

- computational characteristics.

I would like to put forward the following concrete actions:

1. A Forum (under the umbrella of SEMIC.EU) to gather together people, research centres, organisations

2. State of the art. A relevant study to be carried out.

3. Coordination of research efforts at European level

4. Workshop or Conference (possibly in Greece to the honour of Aristotle, founder of his categories)

5. a number of pilot studies

6. a number of well defined projects for pragmatic results

Our ultimate aim is to introduce the Eilenberg - McLane's Category Theory into the Semantic Interoperability meta-level towards the 'standardisation of a communication language' among the relevant European scientific community dealing with open questions and problems of a semantic nature, during the course of communication between living and non-living organisms.

Epilogue

There is still a long way to go before some kind of true Semantic Interoperability will be achieved.

Charalampos Meletis, GR, works within the Ministry of the Presidency and Ministry of Interior in various posts and projects regarding Government IT Strategy, IT standards, eGovernment. He is the Ministry's representative to IDABC, i2010 eGov and European FP6 project regarding Semantic-Gov. In Athens he studied Mathematics and holds a Master in Computer science and a Ph.D. in Cybernetics from the UK. He is the founder (1988) of the Greek Artificial Intelligence Society.