Semiography and Generative Semantics
The IEML Approach

(Information Economy MetaLanguage)

Pierre Lévy et Sylvain Boucher
Canada Research Chair in Collective Intelligence (Ottawa University)

1. Abstract

The **Information Economy MetaLanguage** (IEML) has been developed by Pierre Levy at the *Canada Research Chair in Collective Intelligence* (Ottawa University) to facilitate the representation and automatic processing of sense units belonging to all kinds of ontologies and natural languages. IEML is a regular language that may be used by a finite automaton. IEML can specify a unique semantic address for each distinct concept. A semantic address is a graph of IEML sentences and is represented by a set of algebraic coordinates in the IEML semantic space. This space structure allows the establishment of a correspondence between a set of addresses and a group of computable functions (translations, ranking, logical operations, etc.) selected by different cognitive communities. This correspondance is called an IEML ontology.

2. Semantic addresses

The language offers a set of methods to construct univocal semantic identifiers that may be interpreted in all natural languages and are compatible with semantic Web norms (like RDF and OWL in particular). Any concept may in principle obtain a unique formal representation as a graph of IEML sentences. The sense of a concept is conventionally represented by its semantic address. The "words" composing the IEML phrases are interpreted collaboratively and put in an IEML dictionary available on line.

With the methods used by IEML, it is possible to draw dynamic maps of communities memory. The language may be compared to a *semiography* as the relation between a concept and its semantic address is analogous to the relation between a terrestrial zone and its location on a geographical map.

3. A generative approach

IEML grammar operates according to formal rules: sense units (represented by their semantic identifiers) are constructed by the application of generative rules to a finite set of primitive **elements** and/or to sense units already constructed.

There are **5** primitive elements. They are very general concepts that may be considered as "universals" of the signification process: **A** (actual-effective), **U** (virtual), **S** (sign-signifier), **B** (interpreter-subject) and **T** (thing-referent). The elements **A** and **U** are aspects of an *action* dimension **O**, and the elements **S**, **B**, and **T** are aspects of a *representation* dimension **M**. The cartesian product of **O** and **M** generates four primary regions **OO**, **OM**, **MO**, and **MM** composed of 25 **events** (as an example, see below the semantic field **OM** of primitive acts).

	S	B	T
U	U→S **To know**	U→B **To will**	U→T **To be able**
A	A→S **To say**	A→B **To commit o.s.**	A→T **To do**

Table **OM**

The combinations of the **5** elements follow a rule of asymmetrical semantic synthesis *source→destination* (**So→De**). The asymmetrical combination (**So**, **De**) isolates a semantic field to which is attributed a concept that, as far as possible, is analogous to the sense called to mind by the semantic identifier (**So**, **De**). The concept location follows some other evident rules. For example, the more general concepts must be rooted in the simpler IEML semantic identifiers; the more particular concepts must be rooted in the more syntactically elaborate semantic identifiers.

The **25** "events" (two elements combinations) are reified and combined to create an upper semantic level. This construction procedure is repeated up to **5** times. The last level contains 10 power 23 possible phrases.

A semantic address is structured according to a formal schema **So**urce→**De**stination / **Tr**anslation, the three roles being played by IEML sentences. There is a void translation if the destination is void. With this formal schema (**So**, **De**, **Tr**) it is possible to address any kind of object having a graph structure.

A priori, there are many ways to translate in natural languages an IEML "word" or phrase. So it belongs to communities of experts to decide - through a rational argumentation and consensus – for the sense units in the IEML lexicon.

4. The usefulness of IEML

IEML offers many advantages. We give here the most evident.

Operations on IEML semantic addresses can be programmed: classifications, orderings, logical operations, filterings, semantic form recognitions, topological transformations, simple displacements…

The correspondance between natural concepts and an algebraic structure allows to create, according to search objectives, different methods to compute distance and semantic ranking. Semantic search engines (operating on IEML semantic metadata) can run a wider variety of automatic functions than ordinary search engines.

The IEML translation of natural concepts may give more semantic transparency to the Web. IEML offers a maximal linguistic interoperability to the search engines and could help expand the collective intelligence abilities on the Web.

(N. B. This text is available on the Web at www.ieml.org)