

Semantic Computing in Multimedia

ACM Multimedia Tutorial

Simone Santini^{*}
Universidad Autónoma de Madrid
Madrid, Spain
simone.santini@uam.es

ABSTRACT

This short overview describes the contents of the tutorial *Semantic computing in multimedia*, which was offered to the participants of *ACM Multimedia 2011*.

Given the impossibility of summarizing properly the contents of the tutorial in just two pages, the purpose of this overview is mainly to introduce the reader to the relevant bibliography.

Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems]: methodology; H.1.1 [Systems and Information Theory]: General systems theory

General Terms

Design, Human Factors, Standardization

Keywords

Semantics, multimedia semantics, context, logic

Semantics, as a cognitive computing topic (as opposed, for example, to the formal semantics of programming languages) began within artificial intelligence and then, beginning in the 1980s, faded from public attention following the then general decline of the interest in symbolic artificial intelligence [15]. Those were the heydays of connectionism, and connectionist machines, *ex hypothesi*, would not model semantics explicitly [5]. However, in the last ten years or so there has been a significant resurgence of the technical discourse on semantics, based on the diffused opinion that the large amount of data available today can be properly managed only through a qualitative leap in the processing capabilities of computing machines [2]. A semantic leap, as it is. This is true especially for multimedia data such as video and audio, whose temporal nature (and the consequent imposition of the temporality of the data to the user) makes browsing awkward [14]. Regardless of the correctness, or even the plausibility, of the preeminence given to semantics for the solution of the data access problems (a relation

^{*}This work was supported by the *Ministerio de Educación y Ciencia* under the grant N. MEC TIN2008-06566-C04-02, *Information Retrieval on different media based on multidimensional models: relevance, novelty, personalization and context*.

on which the jury is still out [16, 18]), the relation between meaning and formal systems (for such are computers) has an obvious scientific significance *per se*, and it is a praiseworthy enterprise whose scientific and cultural value can stand on its own.

Semantics is a complex issue, with a history of many centuries and a variety of different points of view (see, for instance, [9, 6] for some bibliography). In order to do serious research on semantics, the computing scientist must be aware of important and complex theoretical questions, and of the solutions and models that the different schools of thought have proposed. The purpose of this tutorial is to provide such background. The idea is to equip researchers with enough technical and philosophical background to allow them to make original and significant contributions to the discipline of semantic computing.

The purpose of this tutorial is not primarily to give the attendance information on standards and programming techniques (which, by now, are fairly well known to the majority of the scientists in the area, and well accessible through a plethora of well written books), but to give them a better view of the larger topics in which semantics is embedded. It is my opinion that learning techniques and standards is the easy part of the job, something that, to quote a popular academic saying, *can be learnt for \$5 worth of late fees at the local library*. The difficult part, the one that needs a face-to-face discussion is, as it is often the case, to know what to do with these standards and techniques, that is, to understand the general theory of semantics and how the different techniques fit in it.

The tutorial is divided in three parts: an introductory one, and two parts dedicated each one to one of two major classes of approaches: the logical and the hermeneutical. The introductory part contained a brief history of semantics and related disciplines, such as logics and semiotics, as well as a brief history of pictorial communication and of the discursive practices that have surrounded the creation of pictures and videos. We shall not go into much depth (this part alone could easily provide material for a Ph.D. if developed fully), but I shall give the attendants enough background to understand the scope and the problems of multimedia communication, and those generated by the deceptively simple concept of multimedia semantics [1, 7].

The second part dealt mainly with logic and with the logical approach to semantics. It covered, roughly, a terrain that goes from Tarski [17] and Carnap [3] to current ontologies [11], with some emphasis on model theory and some foray into partially uncharted waters, such as the use of fuzzy

logic for multimedia modeling. This section, in turn, was divided in two part. The first, and longest, was a technical discussion on the different concepts of semantics that have been used in logic, with special emphasis on aspect of formal semantics and on traditional knowledge representation (including ontologies, the semantic web, and their relation to multimedia). The second part contained a brief excursus on the presuppositions that underlie this work. Traditional logic is a discipline of formal reasoning and never quite dealt with the content (viz. the semantics) of statements. Ever since Aristotle, the systematization of the syllogism operated by the Scholastics, all the way to the axiomatic programme of Russell and the Analytic Philosophy [8], logic has been a science of the *forms* of reasoning, without any reference to the contents of the reasoning activity. If one looks, for example, at the classification of syllogisms, one will find a classification into types of phrases (*Barbara Darii celarent,...*) and a division into figures based only on formal properties of the sentence. Nothing is said of the content of these sentences. The only notable exception is Hegel who, in his *objective* logic, considered that the forms of logic were the basis of metaphysics, that is, of ontology. Alas, every attempt at developing formally Hegel's logic has failed.

To say that the meaning of a document (multimedia or otherwise) can be characterized by a logic theory requires certain assumptions, at the basis of which is the idea that a document has a *content*, independent (more or less) of the linguistic means that are used to express it, and that exists (more or less) intact even if nobody is interpreting the document. We analyzed critically this view of meaning, its plausibility, and the limits of its validity. In particular, we shall consider the notion of *semantic games* as a possible formalization of the Wittgenstenian notion of meaning as use [19].

The incipit of the third part formed a continuum with the non-technical discussion of the second part. While analyzing the presuppositions of the logic approach to semantics, we began to look at alternative views, with a special emphasis on two areas: *hermeneutics* [12] and *structural semantics* [10]. We worked to understand the rôle of the reader in the creation of meaning, the rôle of the discursive practices of media creation, and that of the cultural conventions that drive the way in which media should be interpreted. The study of signification carried out in this way unveiled several important characteristics for the design of semantic systems. Meaning is not an attribute of an image or a video, but something that arises when an artifact is used as part of an *activity*, and only takes meaning in the context of that activity. That is, meaning is created when an artifact is interpreted in a context, and as part of an activity. So, rather than modeling the content of documents, we should model the activities that require access to the images and the context in which these activities take place. We considered several possible solutions for context modeling, highlighting that the variable nature of context makes it imperative to use methods less brittle than logic. Consequently, we studied *soft computing* solutions, starting from the solutions adopted in information retrieval and content based image retrieval [13]. We considered the techniques oriented to a semantic characterization of the data, from latent semantics and self-organizing maps in information retrieval, to semantic classification and cooperative systems in content based image retrieval [4].

The participants to the workshop received a booklet of notes. An improved and continuously updated version of this material is available from the web page of the author: www.ii.uam.es/~ssantini

1. REFERENCES

- [1] Gonzalo Abril. *Análisis crítico de textos visuales*. Madrid:Sintesis, 2007.
- [2] T. Bench-Capon and G. Malcom. Semantics for interoperability: relating ontologies and schemata. Technical Report ULCS-02-029, Dept. of computer science, University of Liverpool, 2000.
- [3] Rudolf Carnap. Empiricism, semantics, and ontology. *Review Internationale de Philosophie*, 4, 1950.
- [4] Scott Deerwester, Susan T. Dumais, George W, Furnas, Thomas K, Landauer, and Richard Harshman. Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6):391–407, 2000.
- [5] Hubert Dreyfus. *What computers still can't do: a critique of artificial reason*. Cambridge, MA:The MIT Press, 1992.
- [6] Umberto Eco. *A theory of semiotics*. Bloomington:Indiana University Press, 1979.
- [7] Norman Fairclough. *Media Discourse*. Bloomsbury, paperback edition, 2009.
- [8] Jerry Fodor. *Concepts*. Oxford:Oxford University Press, 1997.
- [9] Horst Geckeler. *Semántica estructural y teoría del campo léxico*. Madrid:Gredos, 1976. Spanish translation of *Strukturelle Semantik und Wortfeldtheorie* by Marcos Martínez Hernández.
- [10] A. J. Greimas. *Sémantique structurale*. Paris:Larousse, 1966.
- [11] N. Guarino. Formal ontology and information systems. In *Proceedings of FOIS 98, Trento, Italy, 6–8 June*, pages 3–15. Amsterdam:IOS Press, 1998.
- [12] Roy J. Howard. *Three faces of Hermeneutics*. Berkeley:University of California, 1982.
- [13] S. Kaski. Computationally efficient approximation of a probabilistic model for document representation in the WEBSOM full-text analysis method. *Neural Processing letters*, 5(2), 1997.
- [14] Yong Rui, Thomas S. Huang, and Sharad Mehrotra. Constructing table-of-content for videos. *Multimedia Systems*, 7(5):359–368, 1999.
- [15] S. Russell and P. Norvig. *Artificial intelligence, a modern approach*. Upper Saddle River:Prentice Hall, 1995.
- [16] Karen Sparck Jones. What's new about the semantic web? some questions. *ACM SIGIR Forum*, 38(2):18–23, 2004.
- [17] Alfred Tarski. The semantic conception of truth and the foundations of semantics. *Philosophy and phenomenological research*, 4, 1944.
- [18] Yorick Wilks. The semantic web as the apotheosis of annotation, but what are its semantics? In *Twentieth National Conference on Artificial Intelligence (AAAI'05)*, 2005.
- [19] Ludwig Wittgenstein. *Philosophical Investigations*. Prentice Hall, 1973.