

On the Meaning of Meaning

Extended Abstract

Flávio Soares Corrêa da Silva

Laboratory of Interactivity and Digital Entertainment Technology
Department of Computer Science – University of Sao Paulo
Rua do Matão, 1010 – São Paulo, BRAZIL 05504-090 fcs@ime.usp.br

Generally speaking, the term *semantics* refers to the relationships between linguistic objects and their meanings. An object – linguistic or otherwise – can only be meaningful *to* someone or something. More often than not, when we refer to meaning and semantics we consider the interpretation of linguistic objects by human beings.

The World Wide Web (WWW) can be comprised as a dynamic repository of linguistic objects. These objects can be persistent textual data and multimedia files, as well as transitory objects that get through the Internet to connect producers and consumers of data – as occurs for example when people are using instant messaging software or VoIP.

The modalities of communications enabled by the WWW are widely diverse. Communications can be synchronous (such as in instant messaging and VoIP) or asynchronous (such as in electronic mail and discussion groups), one-to-one, one-to-many or many-to-many, personalized or broadcast-like, etc.

Evidently, the WWW would be useless if its contents were deemed meaningless by their users. The WWW is, therefore, fundamentally and inherently meaningful.

The Semantic Web extends the WWW, presumably by adding *semantics* to its contents. It has been the subject of a plethora of articles, projects, consortia and entrepreneurships. Among the most cited articles describing the Semantic Web, we find the one published in the Scientific American magazine (Berners-Lee et al, 2001). Essentially, in that article the authors suggest that the WWW should be extended so that its contents can be interpreted unequivocally by machines as well as humans. Given that the WWW was already meaningful to humans, the proposed extension was so that the Web became also meaningful to machines.

In the present article we analyze the unusual notion of *meaning* that has been employed in the characterization of the Semantic Web. Our view is that a better understanding of the presuppositions and conceptualizations underlying the Semantic Web shall enable its designers and users to explore better the novel resources it brings about.

Indeed, the Semantic Web is not at all about bringing meaning to the WWW – the WWW was already full of meaning, perhaps over flooded with meaning, before the Semantic Web. Actually, the Semantic Web is about *decreasing* the amount of meanings that can be extracted from the contents of the Word Wide Web. It is about reducing the set of potential interpretations associated to each piece of information conveyed by the Word Wide Web to a singleton set. Assuming that the single element of this singleton set is known and agreed upon *a priori* by all users – i.e. producers and consumers of contents for the WWW – it can be informed to software designers and programmers, who then can share this intended meaning with the end users of the contents of the WWW. Software designers and programmers can also assume this

intended meaning when building their pieces of software, whose behavior will then be related with and dependent on this single, pre established interpretation.

Software designed and built based on a pre established interpretation of the data it is going to manipulate can operate based on the explicit representation of data as well as on its implicit intended meaning. Software design and programming is therefore facilitated by the imposed simplification of the intended semantics of the WWW proposed by the Semantic Web initiative. Software modules can become more compact and interoperate more easily.

This has been characterized as making the WWW *machine interpretable*.

There is a price to be paid to make the contents of the WWW capable of direct manipulation by machine. It is precisely the price of agreeing *a priori* on the singleton sets that are going to be assumed as the semantics of the WWW. The users of the Semantic Web must therefore accept the burden of learning the intended meaning of its contents before using it. Any misinterpretation of a piece of data – and any undesired consequence of this misinterpretation – shall be under the responsibility of the user who misinterpreted that piece of data, i.e. interpreted it differently from the software designers and programmers who have built software for the WWW that crunches that specific piece of data.

How authoritarian must be the software designers and programmers, and how submissive must be the users, in order to enable the concept of the Semantic Web? A hint to this answer can be found in the actual notions of *semantics* that have been used in Computer Science, and their applicability in the new context proposed by the Semantic Web.

Through this analysis, we shall conclude that the semantics of the contents of the WWW, already “extended” by the notions of the Semantic Web, are going to be determined by the semantics of the corresponding programs that are going to be processing those contents. The users must be informed, accept and adapt themselves to those semantics, in order to be prepared to use the Semantic Web.

The three classical alternatives for the semantics of programs are (Mitchell, 1996; Corrêa da Silva and Vieira de Melo, 2003):

- *Axiomatic semantics*, according to which the intended meaning of a program can be grasped by comparing it with similar programs. A program means the same as a collection of programs that, given the same input data, generate the same output data. The meaning of a program is conveyed by its synonyms, just as the meaning of a word is conveyed by the synonyms we find for it in a dictionary.
- *Operational semantics*, according to which the intended meaning of a program is determined by the sequence of operations it performs. A program means what it does when it is being processed by a machine.
- *Denotational semantics*, according to which a program is a mathematical mapping between a set of input data and a set of output data. The intended meaning of a program is given by the pairs of input/output sets determined by it.

Based on these characterizations of semantics of programs, we have three alternatives for the semantics of data that is processed by programs:

- Based on the *axiomatic semantics*, the meaning of data is furnished by the programs that process that data. The intended meaning of a piece of data is determined by the set of programs that process it and produce the same results.
- Based on the *operational semantics*, the meaning of data is given by the computations they enable. The intended meaning of a piece of data is conveyed by the operations it can trigger.
- Based on the *denotational semantics*, the meaning of data is given by to what set of equivalent input/output data it belongs. The meaning of a piece of data is conveyed by its synonyms, and what makes two pieces of data synonyms is a program that takes any of them indistinguishably and processes them equivalently.

The Semantic Web, therefore, is the WWW plus a set of programs that process its contents. The users of the WWW must know the intended meaning of the contents of the WWW, employed by the software designers and programmers who build these programs, in order to use and interpret the contents of the WWW appropriately.

The semantics of the Semantic Web shall be defined by the programs that inhabit it. Software production for the WWW is complex, open, highly distributed process which addresses multiple goals. The appropriate study of the semantics of the Semantic Web must relate to the deep and thorough understanding of this process.

Bibliography

Berners-Lee, T., Hendler, J., Lassila, O. *The Semantic Web*. Scientific American, May 2001.

Corrêa da Silva, F. S., Vieira de Melo, A. C. *Princípios de Linguagens de Programação*. Edgard Blucher, 2003.

Mitchell, J. C. *Foundations for Programming Languages*. MIT Press, 1996.