THE SEMANTIC WEB'S PHILOSOPHICAL BACKGROUNDS – GOOGLE SEARLE

If we decide to approach the Internet in more abstract terms, up until today, the Web has been chiefly a social phenomenon. **In its essence, the Web of documents, with all its diversity of content and content-creators, has basically been a network of interconnected documents**, the relations between which were entirely arbitrary, with no principles governing them. It is the website developer on whom form and content depend, and even with the development of social networking and user-generated content (the so-called Web 2.0), the fundamental *structure* of the Web has not changed. From a philosophical point of view, all content on the Web could be limited to syntax alone – and such is the basis on which we can form any models of the present day Internet. Any fundamental regularities that we can perceive on the Internet are fundamentally of human (that is: arbitrary, or contingent) origin.

The Semantic Web is the first proposal and, at the same time, the first actual attempt to bring a new paradigm – a paradigm of highlighting **meaning** as the core building block of the Internet, regardless of the technology or language behind it.

In order to illustrate the difference between the syntax-driven approach as opposed to the semantic paradigm, one of the most famous thought experiments of 20th century philosophy can be invoked: John Searle's Chinese Room experiment. Although it is an argument the main goal of which is to attack the concept of sentient AI, it nonetheless provides a vivid depiction of **the difference between syntax and semantics.**

Searle's thought experiment can be summarized as follows: Let us imagine that a computer program has been created that behaves as if it actually understood Chinese. It can receive input provided in Chinese and provide answers in coherent and grammatically sound manner. It is logically (and technically) possible that such a piece of software could be created, and perhaps even pass the Turing test. Therefore, it could be said that since the machine behaves as if it understood Chinese, then it in fact does understand it.

What Searle questions here is the very notion of understanding. Does the machine actually *understand* the meaning of the words it uses, or is it a mere simulation, despite of the fact that the output provided could not be distinguished from utterances produced by a human

being? To put it in the perspective of our investigation: **how does syntax differ from semantics?**

In order to prove that that it is in fact a mere simulation, Searle asks us to imagine a person that does not know Chinese at all and is isolated from the world in a room in which he receives strings of Chinese characters through a slot in the door. The person then conducts operations on them according to a set of instructions. **What he essentially does is identical to the workings of the software**, the only difference laying in the fact that it is done manually rather than automatically. The output provided by the person in the Chinese Room will also be perfectly sound and appear as if a Chinese-speaking person produced it. However, it is clear that this man has no actual understanding of Chinese, he merely follows rudimentary instructions.

The argument is directed against strong A.I., but in particular it attacks the notion that operations on syntax alone, however convincing their results may be, **cannot imply true understanding.**

The Chinese Room argument (..) (does not) purport to show that machines cannot think—Searle says that brains are machines, and brains think. It is directed at the view that formal computations on symbols can produce thought.

Cole, David, "The Chinese Room Argument", The Stanford Encyclopedia of Philosophy (Winter 2009 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/win2009/entries/chinese-room

Semantic Web research can be ascribed to the broader area of artificial intelligence study. However, it has never been claimed that introducing semantics to the Web would bring it closer to sentient A.I. in any sense – this is still, in essence, procedure-following, and no claims of conscious understanding should be made.

Nevertheless, the Chinese Room argument may be used in order to emphasize the difference between the traditional, syntax-based approach to content on the Web, and how it is dealt with within the Semantic Web. What search engines can accomplish today, is what Searle's operator locked in the Chinese room would do if he was told to look across volumes of incomprehensible texts for sequences of characters – **compare strings with other strings, without any room for reasoning**. The Semantic Web on the other hand introduces vocabularies (ontologies) which ascribe meaning and relationships between items. Within its

paradigm, an object is no longer merely a collection of characters placed in a specific order, it is an entity with particular characteristics and relationships. Without logical operations such as inference, nothing can be done aside from blatant comparison of strings.

According to this new paradigm, the Web's main component is not the document (a formal representation), but the *object with properties*. This means that the fabric of the Web does no longer consist of arbitrarily selected relations (such as hyperlink connections); the objects themselves – with their individual characteristics – form a network. The place of an entity within the Semantic Web is determined by its traits, just as is the case with objects in reality (with dimensions, values, social meaning, etc.). We can talk of the existence of an object, regardless of how we define it – it is *thrown into the system* of other objects with its specific properties. The outcome of this event is not inscribed in the object itself, it is determined by its properties and how they are reacted to by other entitles, just as it is the case with real-world objects.

In this sense, the Semantic Web is much closer to natural language than the traditional, document-based paradigm. **Each element has properties, and its role in the system springs most importantly from its *content*;** location and syntax cease to be essential. The Chinese Room can no longer exhaust the essence of the Web, actual **meaning** becomes decisive.