The semantic web is “an extension of the current web that will allow users to find, share and combine information more easily” (Berners-Lee, 2001, p.34) This logical definition of the semantic web and semantics (meaning instead of mere interpreting) was defined in the equally logic and thought provoking article “The Semantic Web”, which clearly discusses what a semantic web is, how it differs from the normal ‘web’ and what the vital components of it are. The author achieves this by using very simple examples and non-confusing theories that clearly show the basics of semantics; but does not delve any further.

The semantic web will bring structure to meaningful webpage content and have structured information collections and sets of rules that will be used for automated reasoning. Berners-Lee states that the semantic web is a revolutionary thing that will “assist the evolution of human knowledge as a whole.” But perhaps this view on semantics is quite misleading and over enthusiastic. The entire concept is yet still too undefined and developed to make such a statement. For one, the ontology’s or formal document that defines term relations is a very complex matter. The ontology based metadata is a hierarchy of 5 concepts that describe the creation of work; each kind of relation between works occurs at a certain level in the hierarchy. Related works then share data at that level in the hierarchy (Weinstein Peter, 1998). Ontology as metadata schemes provide controlled vocabulary of concepts and explicitly defined machine processable semantics. This entire concept was not discussed by Lee, and as he was discussing the future of semantics he also missed a valuable point; The DAML + OIL layer of semantics; which is still under development. The layer hopes to greatly improve the resource description framework, where meaning is expressed into triplets (subject, verb, objects written using XML (extensible mark-up language), which is a natural way to describe the vast majority of processed machine data) by retaining reification support and extending the RDF concepts which will result in a sounder formal semantics (Hypermedia and the Semantic Web: A Research Agenda, 2005).
Because most of today’s web content is only designed for humans to read, computers would normally have no way to process semantics, “data transmitted across the web is largely throw-away data that looks good, but has little structure” (Hypermedia and the Semantic Web: A Research Agenda, 2005).

The benefit of this more formal ontology scheme is very precise query search results. Berners-Lee also only discussed ontology as a bare-bones subject, when there are in fact many different ontology schemas including a work hierarchy (conception, manifestation, expression, materialization, digitalization, and instance) and the genre structure. Lee failed to mention this great proliferation of formal ontology’s that the semantic web relies upon (Ontology Learning for the Semantic Web, 2001).

However most readings agree on the three basic technologies of the semantic web the Berners-Lee has discussed. The resource description framework (RDF) also contains a universal resource identifier (URI) which identifies the subject and object part of the triples encoded, which forms webs of related information. To sum this up as a practical example used by Lee, multiple databases contain information about people, if a user wants to find a person with a certain zipcode they need to know which fields represent names and zip codes, the RDF specifies this (field 6 in database A) (is a field of type) (zip code) using URI’s instead of phrases (Berners-Lee, 2001, p.34).

The RDF theory has since been advanced on; the RDF-S (schema) defines some pre-defined concepts including subclasses and class hierarchy models (hierarchy models were not clearly defined by Lee) and domain and range restrictions on properties, which generally gave limitations on access and authorization controls.

Another technology is the XML which allows for the creation of hidden tags that are used as notes in WebPages; these allow users to add structure to their documents, although it does not explain it.

The ontology is then the third semantic web component, in which two terms exist. Taxonomies define the classes of objects and relations (an address may be defined as a location type, so city codes may be defined to a certain location), and the inference rules (already discussed) supply the definition with further power (if a city code is then associated with a state code, and an address uses that city code, then that address has that associated state code). However criticisms of these terms describe taxonomies as a hierarchical tree structure (not described by Lee) that is usually too expensive to maintain and create and fails to reflect the language of users (Porter Joshua, 2005).
The solution to taxonomies, or fail this an alternative is a folksonomy, which would let users add tags of information and create navigational links out of these tags to help users find and organise that information later.

Within ontology’s there are still problems with different identifiers, what if two databases use 2 different identifiers for zip code (post code)? As a part of the inference rules equivalence relations (a rule that says this equals this also) solves this dilemma.

Software agents are a crucial aspect of the semantic web; ontology’s are the key to enabling the agent to understand what is on the semantic web and interact with it and other services involving it. Agents use digital signatures which are ‘encrypted blocks of data that computers and agents use to verify the attached information has been provided by a specific trusted source” (Berners-Lee, 2001, p.34). They also works of ‘proofs’ which provide verification and evidence that the search result is correct, and follow a distinct value chain; subassemblies of information that is passed from one agent to another, each one adding value, to construct the final product. Berners-Lee then clearly defines the software agent as the connector to the semantic web, and the impact the agent has on organising and finds meaning in the web.

Metadata is “structured information that describes, explains, locates and makes it easier to retrieve and manage an information resource” (Koyle Karen, 2005, p.160). Interoperability defines multiple systems with different hardware, software, data structures and inferences that exchange data with minimal loss of contact and functionality. Both definitions work well in describing what the semantic web uses and hope to achieve in the electronic world.

However Berners-Lee believes strongly that the semantic web can be implemented in the physical world that the URI’s can point to anything and would use the RDF-S to describe devices such as TV’s and mobile phones. Steps have begun to develop standards for describing functional device capabilities and user preferences, thus “achieving automation with minimal human intervention” (Berners-Lee, 2001, p.34).

Berners-Lee believes the semantic web will “open up the knowledge… of human kind to meaningful analysis by software agents, providing a new class of tools by which we can live, work and learn together.”
His discussion about the physical possibilities of the semantic web as a possible one is a turning point in his discussion. As most of the bare theories laid out have not been fully discussed and defined, he succeeds in explaining the semantic web in a simple fashion, mentioning the 3 technologies that every source agrees with. The only issue found in Lee’s paper was the lack of un-development, and that some areas that address semantics such as hypermedia have not been touched upon. Other than this, he has succeeded in at least defining semantics on the web and the unbelievable possibilities it will hold.
References

Journal Articles


