ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 3, March 2013

The Semantic Web Converting the Current Web Services

Imran Alam
Deptt Of CS
Jamia Hamdard, Delhi

Shoeb Ahad Siddiqui
Deptt Of CSE
Integral University, Lucknow

Nida Khan
Deptt Of CSE
Integral University, Lucknow

Abstract: *In this paper we propose research on how the Semantic Web converting the current web of unstructured documents into a "web of data". The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. The outlined a web of information that provided meaning and structure to information in a way that the existing web did not, which oversees the development of proposed Semantic Web standards. The Semantic Web as a web of data that can be processed directly and indirectly by machines. In the Semantic Web data itself becomes part of the Web and is able to be processed independently of application, platform, or domain. This paper describes how Web services will become more agent-like, and how the envisioned capabilities and uses for the Semantic Web will require implementations in the form of multiagent systems. It also describes how the resultant agent-based Web services will yield unprecedented levels of software robustness.*

Key words: *web of data, multiagent systems, software robustness.*

1. Introduction

The Semantic Web is the extension of the World Wide Web that enables people to share content beyond the boundaries of applications and websites. In the Semantic Web data itself becomes part of the Web and is able to be processed independently of application, platform, or domain. This is in contrast to the World Wide Web as we know it today, which contains virtually boundless information in the form of documents. We can use computers to search for these documents, but they still have to be read and interpreted by humans before any useful information can be extrapolated. Computers can present you with information but can't understand what the information is well enough to display the data that is most relevant in a given circumstance. The Semantic Web, on the other hand, is about having data as well as documents on the Web so that machines can process, transform, assemble, and even act on the data in useful ways. The vision of the Semantic Web is a "web of data" that not only harnesses the seemingly endless amount of data on the World Wide Web, but also connects that information with data in relational databases and other non-interoperable information repositories, for example, EDI systems. Considering that relational databases house the majority of enterprise data today, the ability of Semantic Web technologies to access and process it alongside other data from Web sites, other databases, XML documents, and other systems increases the amount of useful data available exponentially. In addition, relational databases already include a great deal of semantic information. Databases are organized in tables and columns based on the relationships between the data they house, and these relationships reveal the meaning (the semantics) of the data. The Semantic Web is focused on machines. The Web requires a human operator, using computer systems to perform the tasks required to find, search and aggregate its information. It's impossible for a computer to do these tasks without human guidance because Web pages are specifically designed for human readers. The Semantic Web is a project that aims to change that by presenting Web page data in such a way that it is understood by computers, enabling machines to do the searching, aggregating and combining of the Web's information — without a human operator.

2. Semantic Web Background

Working more than 10 years on the Semantic Web's foundations the opportune moment to look at the field's current state and future opportunities. More recently a new technology, Web 2.0 is introduced which is mostly a social revolution in the use of Web technologies, a paradigm shift from the Web as a publishing medium to a medium of interaction and participation. W3C is working on new approaches, such as Gleaning Resource Descriptions from Dialects of Languages (GRDDL) and RDFa, to standardize the linking of structured data with instructions on how to transform or embed data into existing Web resources. The markup of Web services in the DAML family of Semantic Web markup languages enables a wide variety of agent technologies for automated Web service discovery, execution, composition, and interoperation. The realization of the Semantic Web is underway with the development of new AI-inspired content markup languages, such as OIL,3 DAML+OIL (www.daml.org/2000/10/daml-oil), and DAMLL (the last two are members of the DARPA Agent Markup Language (DAML) family of languages). These languages have a well-defined semantics and enable the markup and manipulation of complex taxonomic and logical relations between entities on the Web.

Although many aspects of the Semantic Web are yet to be explored, and much research remains to be done. The interest in semantic web is growing very fast. Semantic web is attracting and will attract the best researchers in the field.

3.Web Extension

The Semantic Web is not a separate entity from the World Wide Web. It is an extension to the Web that adds new data and metadata to existing Web documents, extending those documents into data. This extension of Web documents to data is what will enable the Web to be processed automatically by machines and also manually by humans. It's based on machine-readable information and builds on XML technology's capability to define customized tagging schemes and RDF’s flexible approach to representing data.

4.History

In the early sixties, the concept of the Semantic Network Model was coined by Allan M. Collins, linguist M. Ross Quillian and psychologist Elizabeth F. Loftus in various publications. It extends the network of hyperlinked human-readable web pages by inserting machine-readable metadata about pages and how they are related to each other, enabling automated agents to access the Web more intelligently and perform tasks on behalf of users. The term was coined by Tim Berners-Lee, the inventor of the World Wide Web and director of the World Wide Web Consortium ("W3C"), which oversees the development of proposed Semantic Web standards. He defines the Semantic Web as "a web of data that can be processed directly and indirectly by machines. Many of the technologies proposed by the W3C already existed before they were positioned under the W3C umbrella. These are used in various contexts, particularly those dealing with information that encompasses a limited and defined domain, and where sharing data is a common necessity, such as scientific research or data exchange among businesses.

5.Purpose

The main purpose of the Semantic Web is driving the evolution of the current Web by enabling users to find, share, and combine information more easily. Humans are capable of using the Web to carry out different tasks. However, machines cannot accomplish all of these tasks without human direction, because web pages are designed to be read by people, not machines. The semantic web is a vision of information that can be readily interpreted by machines, so machines can perform more of the tedious work involved in finding, combining, and acting upon information on the web. The Semantic Web is a system that enables machines to understand and respond to complex human requests based on their meaning. This is a challenging task that such an understanding requires that the relevant information sources are semantically structured.

6.Semantic Web Family:

6.1 Uniform Resource Identifier (URI)

A URI is simply a Web identifier: like the strings starting with http or ftp that you often find on the World Wide Web. In fact, the World Wide Web is such a thing: anything that has a URI is considered to be "on the Web". The Semantic Web is generally built on syntaxes which use URIs to represent data, usually in triples based structures: i.e. many triples of URI data that can be held in databases, or interchanged on the World Wide Web using a set of particular syntaxes developed especially for the task. These syntaxes are called "Resource Description Framework" syntaxes.

6.2 Resource Description Framework (RDF)

RDF is a general framework for describing a Web site's metadata, or the information about the information on the site. It provides interoperability between applications that exchange machine-understandable information on the Web. Putting information into RDF files, makes it possible for computer programs ("web spiders") to search, discover, pick up, collect, analyze and process information from the web. The Semantic Web uses RDF to describe web resources. A triple can simply be described as three URIs. A language which utilizes three URIs in such a way is called RDF: the W3C have developed an XML serialization of RDF, the "Syntax" in the RDF Model and Syntax recommendation. RDF XML is considered to be the standard interchange format for RDF on the Semantic Web, although it is not the only format. Once information is in RDF form, it becomes easy to process it, since RDF is a generic format, which already has many parsers. Let’s take an example of XML RDF right now:-

```
 <rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:foaf="http://xmlns.com/0.1/foaf/" >
  <rdf:Description rdf:about="">
       <dc:creator rdf:parseType="Resource">
         <foaf:name>Sean B. Palmer</foaf:name>
       </dc:creator>
       <dc:title>The Semantic Web: An
Introduction</dc:title>
    </rdf:Description>
  </rdf:RDF>
```

6.3 Web Ontology Language (OWL)

OWL is W3C’s specification for creating Semantic Web applications. Building upon RDF and RDFS, OWL defines the types of relationships that can be expressed in RDF using an XML vocabulary to indicate the hierarchies and relationships between different resources. In fact, this is the very definition of “ontology” in the context of the Semantic Web: a schema that formally defines the hierarchies and relationships between different resources. Semantic Web ontologism consists of taxonomy and a set of inference rules from which machines can make logical conclusions.

7. Semantic Web solutions

The Semantic Web takes the solution further. It involves publishing in languages specifically designed for data: Resource Description Framework (RDF), Web Ontology Language (OWL), and Extensible Markup Language (XML). HTML describes documents and the links between them. RDF, OWL, and XML, by contrast, can describe arbitrary things such as people, meetings, or airplane parts. These technologies are combined in order to provide descriptions that supplement or replace the content of Web documents. Thus, content may manifest itself as descriptive data stored in Web-accessible databases, or as markup within documents (particularly, in Extensible HTML (XHTML) interspersed with XML, or, more often, purely in XML, with layout or rendering cues stored separately. The machine readable descriptions enable content managers to add meaning to the content, i.e., to describe the structure of the knowledge we have about that content. In this way, a machine can process knowledge itself, instead of text, using processes similar to human deductive reasoning and inference, thereby obtaining more meaningful results and helping computers to perform automated information gathering and research.

8.Semantic Web Present and Future

It’s important to note that implementation of RDF, OWL, and the Semantic Web as a whole will be a gradual process. Questions about what the Semantic Web is and how it can benefit businesses and individuals are similar to initial confusion about why we needed HTTP and the Web before “WWW” was a staple of our daily vocabulary. But considering how those technologies have proliferated, it’s likely that the Semantic Web vision is one that will be realized, even if it’s on a small scale initially. It’s also important to note that, similar to current Web services implementations, the Semantic Web may initially be restricted to intranet and extranet applications until questions about information security can be sufficiently addressed. The true impact of the Semantic Web will not be known for quite some time, but its potential is staggering. Some Semantic Web proponents have asserted that it will lead to the evolution of human knowledge itself by allowing people - for the first time - to quickly filter and synergize the massive amounts of data that exist in the world in a relevant, productive way. The Semantic Web is a mesh of information linked up in such a way as to be easily process able by machines, on a global scale.

9.Proposed Architecture

The proposed architecture mainly depends on sorting algorithm. Suppose there are a number of sorting algorithms available. Each might have possibly errors, weaknesses and strengths. Each process might be faster or slower but be able to sort strings as well as integers. The main thing is that how can the algorithms be combined so that the strengths of each are exploited and the weaknesses of each are compensated. With the help of this approach we hypothesize that the end result is an agenizing of each algorithm. The first approach is proposed and used by us is preprocessing algorithm which receives data and sorting them and would then choose the best algorithm to perform the sorting. The characteristics of each module have to be encoded into the central unit. The best result depends upon the known facts about each of the module in the central unit. With the help of simplistic algorithm, the central unit determining the best result. Of course the various difficulties and challenges might be arising in this approach. The different problems with this approach are (i) it is difficult to add this approach as a new algorithm in the existing algorithm (ii) this approach may have weaknesses or mistakes (iii) it is possible that only one module executes at a time, if it happens then there is low CPU usage when process is completed. In this case the most CPU resources are wasted.

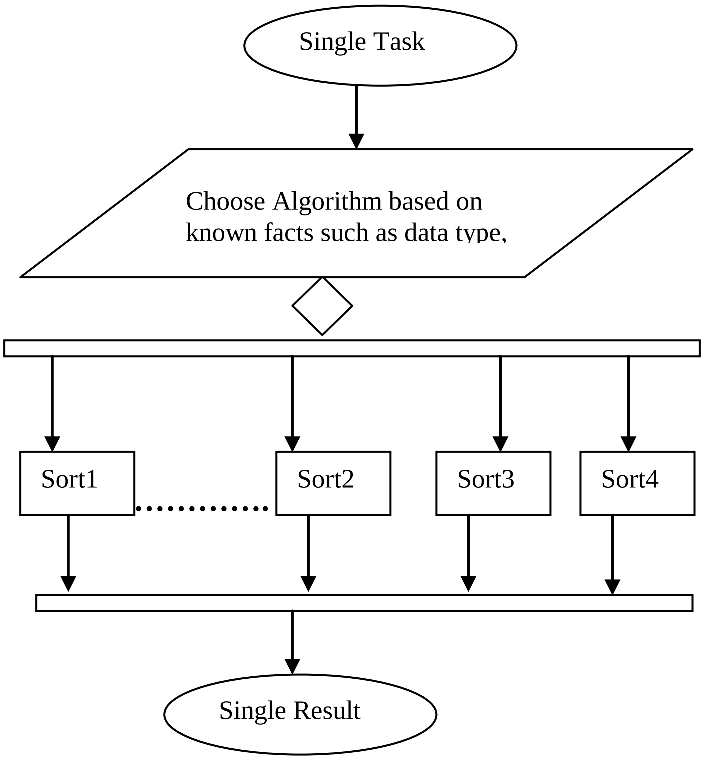

Fig 1: Architecture of a sorting algorithm where a preprocessing algorithm chooses which sorting algorithm will execute.

Due to some lackness in the preprocessing algorithm an improved and second approach post processing algorithm is introduced that receives the results of all sorting algorithms and chooses the best result to be the output. Results have to be compared and select the best. This improved approach is also centralized and suffers from a waste of CPU resources, because all algorithms work on the data. However, the comparison among them produces better results.

The next and third approach is the combination of the preprocessing and post processing centralized systems could also be used. Since the characteristics of each module are known in the central unit. So a subgroup could be selected to perform the desired task based on known factors namely time complexity, space constraints and speed. The results of the subgroup will be compared to determine the best result as above. If every time the certain modules will be selected then the same set of circumstances comes up. A better approach, the developing a conventional algorithm would eliminate the need for a central intelligent filtering unit

ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 3, March 2013

The next and fourth approach is a distributed solution. In this approach the algorithms jointly decide which one(s) should perform the sorting, and if there is more than one qualified algorithm, they jointly decide on the best result. Since conventional algorithms do not have a distributed decision-making ability, so we firstly investigated a generic capability that provides decision making ability to conventional algorithms. If a generic capability can be added to an algorithm, then the conventional algorithms enable to participate in a distributed decision. We discovered that the result has the characteristics of a sorting agent.

Proceeding in the same manner we collected a number of sorting algorithms and each algorithm written by different people and it's having different characteristics such as input data type, time complexity, space complexity and output data type.

On the basis of our experiments we converted each algorithm into a sorting agent and each sorting agent is composed of a sorting algorithm and a wrapper for that algorithm. The wrapper is associated with the algorithm but does not know about the inner workings of the algorithm. It has knowledge only about the characteristics of known facts of its algorithm such as the data type(s) it can sort, its time complexity, its space complexity and the data type it produces

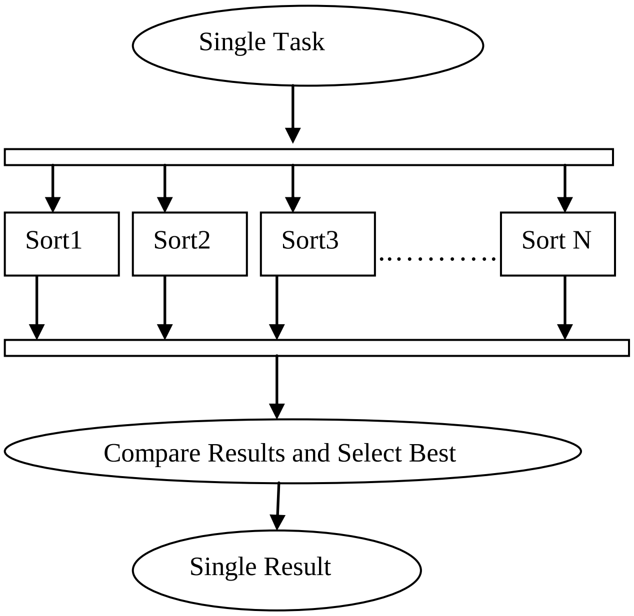

Fig 2: Architecture of a sorting algorithm where a post processing algorithm chooses one result to be the output.

After that the system sends data to be sorted to all the sorting agents. In a group the sorting agents enable to fulfill their responsibility to sorting of the data and they should be able to do this better than any one of them alone. The sorting process of the sorting agents depends upon the data type, its speed and its own knowledge of what types it can sort and how it can be sorted. After receiving data to be sorted each agent determines whether or not it can sort it successfully and if the agent can do then it sends a message to every other agent identifying its intention along with a measure of performance for its algorithm, which is based on time and space complexity. The decision of agents to be chosen among those that have capability to make the sorting data in a distributed manner. Upon receiving the estimates from the other agents each agent compares its own performance measure against those received in the messages. If the agent exhibits the best performance measure, it will run its algorithm and send the results back to the system. Otherwise it will do nothing. The results showed that the composition of sorting algorithms performed better than any individual algorithm. On the basis of different observations any individual algorithm may not solve the problem of semantic web.

10. Conclusion and Future Work

It is difficult to predict the future without uncertainty. Creating, developing and producing strong and healthy software is much more difficult. Implementing and distributing it across the Web is much more complex and complicated. The recommended approach would be helpful for the developers to construct software systems:

1. The coding of this algorithm should be error free. The main attention of developers should be on the development of the algorithm.
2. No any one algorithm may solve the problem of semantic web. Agents (Algorithms) are easier to reuse and easier to add to an existing system because agents are interact with an arbitrary number of other agents.
3. Without a centralized controller unit Agents will need to understand how to detect and correct inconsistencies in each other's behavior.

4. In the field of software development several people might be write an algorithm and the best algorithm will be selected.
5. Agent organizational identifications need to be developed to take full advantage of redundancy.

Under the suggested approaches here, all algorithms would be selected depending upon the developers of the web services. The developers can be used to together for a single company. But the main aim is that how the system tie together to solve the problem.

11. References

1. J. Hendler, "Agents and the Semantic Web," *IEEE Intelligent Systems*, vol. 16, no. 2, Mar./Apr. 2001
2. T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific Am.*, May 2001.
3. The Semantic Web: A Guide to the Future of XML, Web Services and Knowledge Management, Michael C. Daconta, Leo J. Obrst, and Kevin T. Smith, John Wiley & Sons Inc (2003).
4. Intelligent Information Integration for the Semantic Web, Ubo Visser, Springer (2004).
5. Semantic Web Services, Processes and Applications, Jorge Cardoso and Amit Sheth (Editors), Springer (2006).
6. Berners-Lee, Tim, James Hendler, and Ora Lassila: "The SemanticWeb," *Scientific American*, vol. 284, no. 5, May 2001.
7. Huhns, Michael N. and Vance T. Holderfield: "Robust Software," *IEEE Internet Computing*, vol. 6, no. 2, March-April 2002.
8. A Developer's Guide to the Semantic Web, Liyang Yu, Springer-Verlag (2011).
9. Foundations of the Semantic Web, Rajendra Akerkar, Oxford: Alpha Science, (2009).
10. http://www.w3schools.com/default.asp
11. http://semanticweb.org