Object-Process Methodology as a basis for the Visual Semantic Web

Dov Dori

Technion, Israel Institute of Technology, Haifa 32000, Israel

dori@ie.technion.ac.il, *and*

Massachusetts Institute of Technology, Cambridge, MA 02139, USA

dori@mit.edu

Abstract

The Visual Semantic Web (VSW) is a new paradigm for enhancing the current Semantic Web technology. VSW is based on Object-Process Methodology (OPM), which enables modeling of systems in a single graphic and textual model. VSW provides for representation of knowledge over the Web in a unified way that caters to human perceptions while also being machine-processable. The advantages of the VSW approach include graphic-text knowledge representation, visual navigability, semantic sentence interpretation, specification of system dynamics, and complexity management. Arguing against the claim that humans and machines need to look at different knowledge representation formats, we present the principles and basics of OPM and RDF. Using a running example, we present VSW as an extension of OPM, where the basic idea is to express knowledge visually and via a subset of natural language, such that the two representations are strictly equivalent. Both the graphics and the text are intuitive, yet formal, so they are not only understandable to humans, but are also amenable to mechanical processing just like XML.

1. Introduction: The Human-Machine Language Orientation Dilemma

A major assumption underlying the development of the Semantic Web is that humans and machines must each use a different format of knowledge representation. The first sentence of the introduction of the Resource Description Framework (RDF) [7] reads: "*The World Wide Web was originally built for human consumption, and although everything on it is* ***machine-readable****, this data is not* ***machine-understandable****"* (emphasis in source).

The implicit, common wisdom assumption, according to which human-readable and machine-readable formats are bound to be different, is also at the basis of OWL, the Web Ontology Language [11]. OWL is intended to provide a language that can be used to describe the classes and relations between them that are inherent in Web documents and applications. The introduction to OWL [11] reads: "*The World Wide Web, as it is currently constituted, resembles a poorly mapped geography. ... In order to map this terrain more precisely, computational agents require machine-readable descriptions of the content and capabilities of web accessible resources. These descriptions* ***must be in addition*** *to the human-readable versions of that information.*"

This work questions the assertion (not emphasized in the source), according to which machine-readable descriptions ***must*** be added on top of the human readable ones. Rather, this work proposes the Visual Semantic Web paradigm, in which the human and machine representations of knowledge are effectively identical, enabling humans to benefit from the advantages of a dual, text-graphic knowledge representation.

2. Human vs. Machine Understanding and Language Readability

The Resource Description Framework (RDF; [2], [7]) aims at making the knowledge resources that are available on the Web amenable to machine interpretation, compilation, and other types of processing, by imposing some structure on the pieces of knowledge. RDF provides a basis for a number of emerging initiatives, such as the Dublin Core Metadata Initiative [6], an open forum engaged in the development of interoperable online metadata standards. One must, however, bear in mind, that machines are never going to "understand" knowledge the way humans do. At best, they can exhibit treatment of this knowledge such that for humans it seems as if they understand it. Indeed, as the Cyc project [3] has demonstrated, analysis of unconstrained natural language is way too difficult. Since a combination of graphics and text is highly effective as a knowledge modeling language [4], this paradigm has been the basis for the design of the Visual Semantic Web.

Based on Object-Process Methodology (OPM), [4] it is the objective of this paper to propose that OPM, with the extensions required to implement the Visual Semantic Web, be adopted as a human-understandable layer on top of RDF [2], [7] as a means to specify knowledge over the Web.

3. RDF Basics and Example

RDF is a model for representing named properties and property values [7]. RDF properties may be thought of as attributes of resources and, in this sense, correspond to traditional attribute-value pairs. RDF properties represent relationships between resources, and RDF schemas, which are instances of RDF data models, are Entity-Relationship (ER) diagrams. The basic RDF data model consists of three object types: *Resource*, *Property*, and *Statement*. Statement is a specific resource, which consists of a *subject*, together with a named property—the *predicate*, plus the value of that property for that resource—the *object*. The object of a statement (i.e., the property value) can be another resource, a literal, i.e., a resource (specified by a URI), a simple string, or another primitive datatype defined by XML. Based on [2], consider the sentence:

Ora Lassila is the creator of the resource http://www.w3.org/Home/Lassila.

Translated to RDF format, this sentence can be interpreted as having the subject (resource) http://www.w3.org/Home/Lassila, the predicate (property) creator, and the object (literal) "Ora Lassila". As noted, RDF uses directed graphs to specify these graphically, where subjects and objects are nodes, and predicates are labels along the edges. Edges are always directed from a subject to an object, as in Figure 1. A resource node in the graph is drawn as an oval (ellipse), while a literal node is drawn as a rectangle.

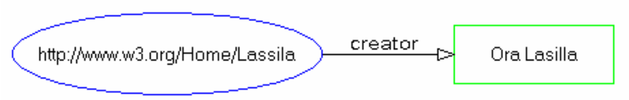

Figure 1. A simple RDF example from [7]

The graph in Figure 1 is to be interpreted as "*http://www.w3.org/Home/Lassila has creator Ora Lassila*", and in general "*<subject> HAS <predicate> <object>*". Applying the XML-based RDF syntax to this graph, specified in [7], one gets the following RDF/XML script:

```
<rdf:RDF>
  <rdf:Description
about="http://www.w3.org/Home/Lassila">
    <s:Creator>Ora Lassila</s:Creator>
  </rdf:Description>
</rdf:RDF>
```

The graph in Figure 1 and the XML script above are not equivalent in their information content, as the XML script contains the XML tags `<rdf:RDF>` and `<rdf:Description>`, which are namespace definitions [1] that are not shown in Figure 1. Since RDF is machine-oriented, it does not emphasize the importance of the graphic knowledge representation, which is provided only for human consumption.

4. An OPM-based Visual Semantic Web alternative

The Visual Semantic Web (VSW) alternative to the RDF/XML knowledge representation takes advantage of the integrated graphic-text formal yet intuitive infrastructure that OPM provides. Figure 2 is a Visual Semantic Web specification (VSW spec), which expresses the example in Figure 1 as an Object-Process Diagram (OPD) and an Object-Process Language (OPL) text. The OPD contains two object instances: **Ora Lasilla** and WWW.w3.org/Home/Lassila. To conform to OMG UML 1.4 [10], object names (i.e., instances of object classes), are underlined in this OPD, as in UML object diagrams.

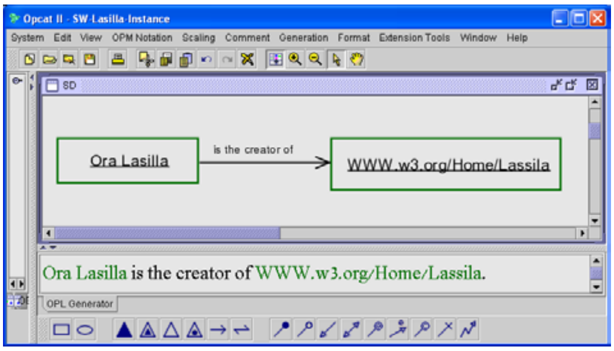

Figure 2. The example in Figure 1 expressed as a Visual Semantic Web specification (VSW spec), consisting of an Object-Process Diagram (OPD) at the top window and its corresponding, automatically-generated Object-Process Language (OPL) sentence at the bottom window.

A tagged structural link, depicted as an open arrow pointing from the **Person** to the **URI**, expresses the nature of the relation between these two objects. The tag is the text recorded along the structural link. The value of this tag is '**is the creator of**'. The value is a phrase, i.e., a collection of one or more words (separated by spaces, as in natural languages), such that when the name of the source object, **Ora Lasilla** (an instance of the class **Person**) is concatenated with the tag value (i.e., the phrase) '**is the creator of**' followed by the name (value) of the **URL**, one automatically gets the following OPL sentence.

Ora Lasilla is the creator of WWW.w3.org/Home/Lassila.

This OPL sentence is also generated automatically by OPCAT and recorded at the bottom of the OPD in Figure 2. The automatic generation of the OPL sentence in this simple case was done by concatenating the name of the object at the source of the tagged structural link, **Ora**

Lasilla, with the text string of the structural link's tag, **is the creator of**, with the name of the destination object, WWW.w3.org/Home/Lassila.

5. The VSW Schema: adding class information

The lines under the two objects in Figure 2 denote the fact that these are instances, not classes. The class information is still missing in this OPD.

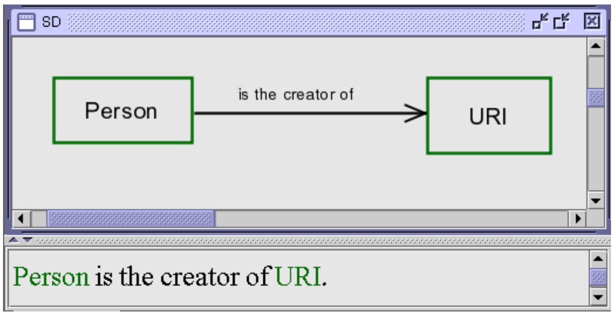

Figure 3. A VSW schema showing the class OPD along with its corresponding OPL sentence, to which the VSW spec in Figure 2 conforms.

Figure 3 shows a Visual Semantic Web (VSW) schema. A *VSW schema* is an OPM model (an OPD-OPL combination) that contains class information. Note that the VSW schema follows the OPM text-graphic equivalence principle, according to which the OPD and the OPL sentence are completely equivalent and therefore reconstructible from each other.

Each VSW spec conforms to a VSW schema. Thus, the VSW schema in Figure 2 conforms to the VSW spec in Figure 3. This VSW schema can be thought of as a template that expresses a rule. In our example, the rule stipulates that the source, which in RDF schema terminology is termed the *domain*, of the relation (*predicate*) '**is the creator of**' is an object that belongs to the class **Person**, and that the destination (*range*) of that relation is an object that belongs to the class **URI**.

Having established the **Person-URI** VSW schema, we can now use it to add the object instance for each of the two classes. This is done in the instantiated schema shown in Figure 4, where the VSW schema of Figure 3 and the VSW spec of Figure 2 are combined. The combination uses the OPM classification-instantiation relation, which is denoted as a bulleted triangle whose tip is linked to the class and whose base is linked to the instance. Note that the instances **Ora Lasilla** and WWW.w3.org/Home/Lasilla need not be underlined here to denote that they are instances. The underlining of the instance names is only mandatory if the class information is not present in the OPD, but here this is indicated by the classification-instantiation links from the classes to the respective instances.

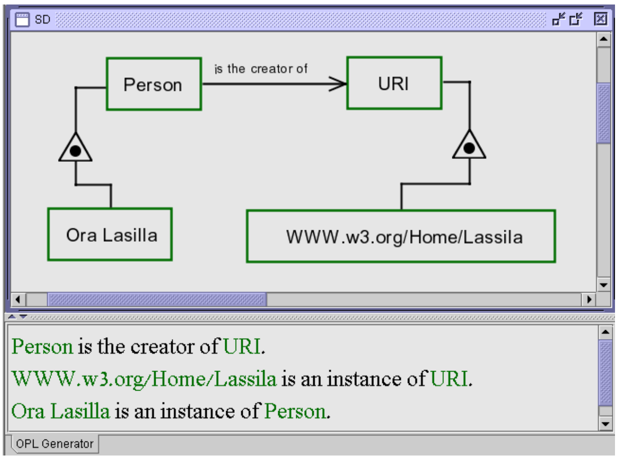

Figure 4. The instantiated VSW schema generated by adding the instance specification of Figure 2 to the class information in the VSW schema in Figure 3.

The automatically generated OPL paragraph of the OPD in Figure 4 (shown also in the OPL window at the bottom of the figure) is:

Person is the creator of URI.
WWW.w3.org/Home/Lassila is an instance of **URI.**
Ora Lasilla is an instance of **Person.**

Note that predicates (such as **is the creator of**) do not have explicit instance names that are distinct from their class names. Thus, for example, we use the same predicate in the OPL sentence "**Mark Twain is the creator of Huckleberry Finn.**" as in "**Ora Lasilla is the creator of** WWW.w3.org/Home/Lassila." The tagged structural relation '**is the creator of**' from the class **Person** to the class **URI** is inherited to their respective instances, so there is an implicit tagged structural relation with the same tag, '**is the creator of**', from **Ora Lasilla**, an instance of the class **Person**, to WWW.w3.org/Home/Lassila, an instance of the class **URI**.

The instantiated VSW schema in Figure 4 has a couple of drawbacks: First, it is space-consuming, and second, it requires the reader to realize the existence of the implicit tagged structural relation. These two problems are solved in the compact version of the instantiated VSW schema of Figure 4, shown in Figure 5.

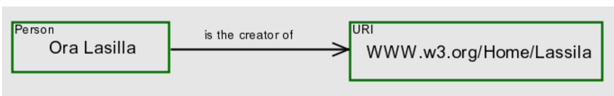

Figure 5. A compact version of the instantiated VSW schema in Figure 4

The OPL paragraph that corresponds to the OPD in Figure 5 is also more compact than the three-sentence OPL paragraph of Figure 4, as it consists of just one sentence:

The **Person Ora Lasilla is the creator of** the **URI** WWW.w3.org/Home/Lassila.

This sentence combines the OPL schema sentence from Figure 2, which is "**Person is the creator of URI.**" with the OPL instance sentence Figure 3, which is "**Ora Lasilla is the creator of** WWW.w3.org/Home/Lassila." In the new OPL sentence, which reflects both the classes and the instances, we added the classes of both **Ora Lasilla**, which is **Person**, and of WWW.w3.org/Home/Lassila, which is **URI**. **Ora Lasilla** is classified as belonging to the class **Person** by preceding the name of the instance by the reserved word "The" followed by the class name **Person**. Likewise, the string WWW.w3.org/Home/Lassila was classified as belonging to the class **URI** by preceding the value of the string by the reserved word the followed by the class name **URI**. The corresponding quoted sentence, used for machine consumption, in which each component of the sentence is surrounded with single quotes, is:

```
The 'Person' 'Ora Lasilla' 'is the creator of' the 'URI' 'WWW.w3.org/Home/Lassila'.
```

Theoretically, lacking quotes, the above OPL sentence contains a potential ambiguity for the human reader, who might think that the class name is '**Person Ora**' while the instance of that class is '**Lasilla**', because the words '**Person**' and **Ora** both start with a capital letter. However, even in this case, a quick look at the corresponding OPD clarifies that '**Person**' and not '**Person Ora**' is the class name.

The OPM text-graphics equivalence principle mandates that any piece of information contained in the OPL paragraph be represented in the corresponding OPD, and vice versa. This principle, which makes the OPD and its OPL paragraph fully equivalent in terms of information content, is followed in the OPM specifications of both Figure 4 and Figure 5.

6. Adding attributes

Continuing with the example from [7], for specifications that are more complex, a compound resource can be created, as the following sentence and the corresponding graph in Figure 6 demonstrate:

"The individual referred to by employee id 85740 is named Ora Lassila and has the email address lassila@w3.org. The resource http://www.w3.org/Home/Lassila was created by this individual."

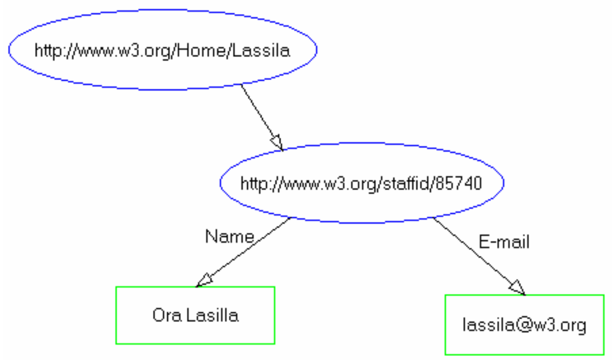

Figure 6. An identified property with structured value [7]

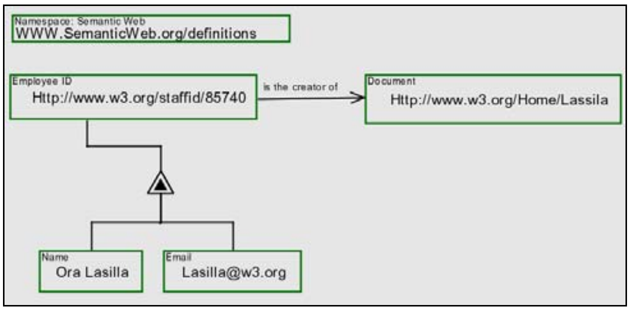

Figure 7. The OPD corresponding to the graph in Figure 6

The OPD in Figure 7 and its corresponding OPL paragraph below correspond to the graph in Figure 6.

The default namespace **Semantic Web** is at WWW.SemanticWeb.org/definitions. The **Employee ID** WWW.w3.org/staffid/85740 **is the creator of** the **Document** WWW.w3.org/Home/Lassila. The **Employee ID** WWW.w3.org/staffid/85740 exhibits the **Name Ora Lasilla** and the **Email** Lasilla@w3.org.

The OPL reserved word exhibits expresses the exhibition-characterization relation (the relation between a class and its attributes, symbolized by a black-in-white triangle) from The **Employee ID** Http://www.w3.org/ staffid/85740 to the **Name Ora Lasilla** and to the **Email** Lasilla@w3.org.

A better representation of the information presented in the OPD in Figure 7 is shown in the OPD of Figure 8. The **Employee ID** is now an attribute of the **Person** rather than the other way around. That this is a better way of modeling is clearly seen when we compare the OPL paragraph below, which corresponds to the OPD in Figure 8, to the previous OPL paragraph, which corresponds to the OPD in Figure 7.

The default namespace **Semantic Web** is at WWW.SemanticWeb.org/definitions. The **Person Ora Lasilla is the creator of** the **Document** WWW.w3.org/Home/Lassila. The **Person Ora Lasilla** exhibits the **Employee ID** WWW.w3.org/staffid/85740 and the **Email** Lasilla@w3.org.

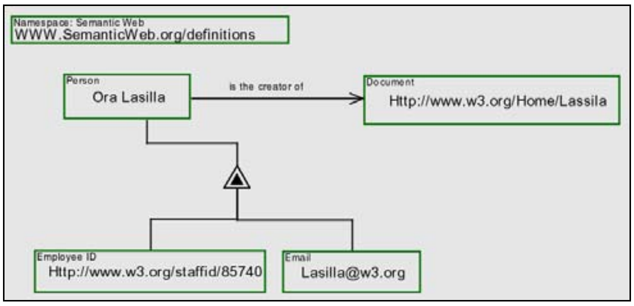

Figure 8. A better representation of the information presented in the OPD in Figure 7

7. Summary and Future Work

The Visual Semantic Web (VSW) paradigm proposes to unify human and machine representations of knowledge. The foundation for this unification is Object-Process Methodology (OPM), which advocates the integration of a system's structure and behavior is a single bimodal graphic and textual model. Using a simple example, the paper has presented the principles and outline of implementation for the VSW. Like OPM, the VSW model enables the representation of static and dynamic knowledge using a combination of Object-Process Language (OPL), a subset of English, and Object-Process Diagrams (OPDs), an equivalent visual formalism. The advantages of this approach include graphic-text knowledge representation, visual navigability, semantic sentence interpretation, specification of system dynamics, and complexity management. We plan to augment OPCAT 2 to handle the extensions presented in this work and provide a Web-based Visual Semantic Web navigation tool.

8. References

[1] Bray, T., Hollander, D. and Layman, A. (1999). Namespaces in XML, World Wide Web Consortium Recommendation, 14 January, 1999. http://www.w3.org/TR/REC-xml-names

[2] Brickley, D. and Guha R.V. (2002). RDF Vocabulary Description Language 1.0: RDF Schema, W3C work in progress draft, 30 April, 2002. http://www.w3.org/TR/rdf-schema

[3] Cyc. OpenCyc, org. http://www.opencyc.org/, 2002.

[4] Dori, D. (2002) *Object-Process Methodology - A Holistic Systems Paradigm*, Springer Verlag, Berlin, Heidelberg, New York.

[5] Dori, D. Reinhartz-Berger, I. and Sturm A. OPCAT – A Bimodal CASE Tool for Object-Process Based System Development. Proc. IEEE/ACM 5th International Conference on Enterprise Information Systems (ICEIS 2003), École Supérieure d' Électronique de l' Ouest, Angers, France, 23-26 pp. 286-291, April, 2003.

[6] Dublin Core Metadata Initiative. http://www.dublincore.org/, 2002.

[7] Lassila, O. and Swick, R. Resource Description Framework (RDF) Model and Syntax Specification. W3C Recommendation, 22 February, 1999. http://www.w3.org/TR/REC-rdf-syntax

[8] Mayer, R.E. *Multimedia Learning*. Cambridge University Press, NewYork, NY, 2001.

[9] OMG UML1.4 (2001). Object Management Group, Unified Modeling Language Version 1.4, September 2001.

[10] OMG XMI (2002). Object Management Group, XML Metadata Interchange (XMI) Specification Version 1.2, January 2002. http://cgi.omg.org/docs/formal/02-01-01.pdf

[11] Smith, M. K., McGuinness, D. Volz, R., and Welty, C. (2002). Web Ontology Language (OWL) Guide Version 1.0. W3C Working Draft, 4 November, 2002. http://www.w3.org/TR/2002/WD-owl-guide-20021104/