Learning Objects on the Semantic Web

Permanand Mohan
Department of Mathematics and Computer Science
The University of the West Indies
St. Augustine
Trinidad and Tobago
(pmohan@tstt.net.tt)

Christopher Brooks
ARIES Laboratory
Department of Computer Science
University of Saskatchewan,
Saskatoon, Saskatchewan,
Canada
(cab938@mail.usask.ca)

Abstract

An important issue in reusing learning objects on the Semantic Web is the development of appropriate technology to facilitate the discovery and reuse of learning objects stored in global and local repositories. Another issue is the development of ontologies for marking up the structure of learning objects and ascribing pedagogical meaning to them so that they can be understandable by machines. A third issue is making learning objects smarter so that they can perform a more meaningful role on the Semantic Web. This paper discusses these and other issues as they affect the exploitation of learning objects on the Semantic Web.

1. Introduction

One of the more recent developments with the Web is an activity known as the Semantic Web. The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation [1, p.35]. Two important technologies for developing the Semantic Web are XML and RDF. XML allows users to add arbitrary structure to documents without saying what these structures mean. RDF allows meaning to be specified between objects on the Web and was intentionally designed as a metadata modeling language.

A third important aspect of the Semantic Web is a set of ontologies. An ontology is a *specification of a conceptualization* [2]. It describes the concepts and relationships of some phenomenon in the world. By using well-defined ontologies on the Web, it is possible for computers to meaningfully process data since there is a common understanding of terms used and the relationships between these terms.

The Semantic Web is concerned about the meaning of all kinds of information on the Web. Learning objects comprise an important subset of this information. A learning object is a digital learning resource that facilitates a single learning objective and which may be reused in a different context. In recent years, the concept of a learning object has received considerable attention in e-learning. It can be very expensive and time-consuming to develop the content for an e-learning course. Being able to reuse learning objects created by others reduces the time and cost to develop learning materials. The learning materials may even be of a higher quality than if developed from scratch, similar to well-designed and well-tested software components.

In order to reuse content from one system to another, it is important for learning objects to be standardized. To this end, there have been a number of global efforts to develop standards, specifications, and reference models for learning objects. Recently, the IEEE 1484.12.1-2002 Standard for Learning Object Metadata (LOM) [ltsc.ieee.org] was released, the first accredited standard for learning technology. The LOM Standard uses nine categories of XML data elements to describe a learning object: General, LifeCycle, Meta-Metadata, Technical, Educational, Rights, Relation, Annotation, and Classification. It is intended that LOM metadata will simplify the discovery, management, and exchange of learning objects over the Web. Another important specification is the Content Packaging Specification from the IMS Global Consortium [www.imsglobal.org], which is a specification for packaging groups of learning objects so that they can be reused in different learning content management systems (LCMSs).

Learning objects developed and stored in many different places on the Web have a tremendous potential to benefit e-learning in particular and education in general. However, there are numerous technical issues that must be dealt with before learning objects can be effectively reused from one situation to the next. These issues are highlighted in this paper. Section 2 discusses the different contexts in which reuse of learning objects is likely to occur and presents our vision for learning objects on the Semantic Web. Section 3 describes the information and technical requirements to achieve our vision. Two key requirements are repositories for storing learning objects and shareable ontologies to describe the structure of learning objects as well as the conceptualization of a domain. These requirements are discussed in Sections 4 and 5, respectively. Section 6 compares the development of learning object technology with the rise of distributed computer systems. Section 7 discusses the need for smarter learning objects to support their enhanced role on the Semantic Web. It also describes how we intend to make learning objects smarter by implementing them as object-oriented learning objects.

2. Vision for Reusable Learning Objects

There are two main types of players in the learning object economy: producers and consumers. Producers of learning objects use various Web-design tools and other software to produce different kinds of learning objects. Consumers of learning objects use learning objects created by others (or themselves) to develop new content packages. Producers make their learning objects available by placing them in different kinds of repositories accessible from the Internet. Typically, consumers are expected to search these repositories using metadata such as that defined in the LOM Standard. There may be copyright and payment issues associated with the reuse of learning objects; however, these are beyond the scope of the paper.

Current development efforts with learning objects are mostly concerned about metadata and content packaging aspects. There has not been any significant work done so far in automating the discovery and packaging of learning objects based on variables such as learning objectives and learning outcomes. There has also not been a significant amount of work done in personalizing e-learning based on learning objects developed and stored at arbitrary locations on the Internet. This is largely because learning objects are a relatively new phenomenon. Automating these processes is also a knowledge-intensive activity likely to require the application of artificial intelligence techniques such as knowledge representation and reasoning.

Increasingly, researchers in adaptive hypermedia, Web-based education systems, and intelligent tutoring systems (ITSs) are turning their attention to personalizing instruction on the Web using learning objects [3]. However, many of these efforts have been based on using learning material developed specifically for the course at hand, and so avoids the problem of assembling course packages out of arbitrary learning objects located on the Web.

Given the vision of the Semantic Web, we believe that learning objects have much greater potential than what is commonly suggested. Consider an instructor or course designer developing a course. The instructor should be able to map out a set of concepts in the domain and the set of learning outcomes that are desired. The instructor should then be able to give this information to a learning object search agent that searches the Web and returns a pool of learning objects that would be appropriate, with alternatives where necessary for imparting concepts in the domain.

Our vision for learning objects on the Semantic Web is that the learning objects themselves should play a more meaningful role in the search process, and should be able to interact intelligently with an LCMS to provide instruction on the Web. If a learning object is able to determine its suitability for an instructional situation, then search agents on the Semantic Web are able to perform more sophisticated searches for learning objects, resulting in pools of learning objects that are likely to achieve the instructional goals. Since all other learning objects on the Web are effectively ignored, the vision for learning objects on the Semantic Web would be realized.

We do not expect learning objects and search agents to independently decide what is suitable for an instructional situation. It is likely that the instructor or course designer will have to manually examine the pool of learning objects returned, to fine-tune their integration into the course. However, the time taken to search for learning objects and package them into a course would have been significantly reduced since the agent only harvested relevant learning objects based on the criteria established by the instructional designer.

3. Requirements to Support Vision

Given the context in which learning objects can be reused on the Semantic Web, it is important to develop the technologies to make it happen. First of all, learning objects must be made available to potential consumers on the Web. This can be achieved by storing them in various kinds of repositories that matches the mode of production of learning object producers. These repositories must provide query services that facilitate the search and retrieval of learning objects.

Secondly, learning objects stored in repositories must also provide semantically rich information to facilitate their discovery and reuse on the Web. To support this requirement, ontologies must be developed for specifying domain concepts and the structure of learning objects. In addition, learning objects must provide enough pedagogical information to enable personalization during an instructional interaction. This information will specify the kinds of cognitive activities in which learners are engaged, and the teaching and learning strategies used in the learning object to impart the concepts of the domain.

Finally, techniques and tools to support the production and reuse of learning objects on the Web must be developed. The next two sections discuss the issue of learning object repositories and the need for shareable ontologies.

4. Repositories of Learning Objects

Global repositories of learning objects are increasingly becoming available on the Internet. These include TeleCampus [telecampus.edu], the Campus Alberta Repository of Learning Objects (CAREO) [www.careo.org], and the Multimedia Educational Resource for Learning and Online Teaching (MERLOT) [www.merlot.org]. Global repositories usually maintain links to learning objects stored elsewhere on the Web, though some of them physically store the learning objects. Several global repositories are using the LOM standard for cataloguing learning objects. However, given the diverse origins of learning objects and the possibility that different kinds of metadata may have been used to catalog the learning objects (e.g., Dublin Core and CanCore – a more usable subset of the LOM [4]), a more flexible approach to specifying metadata is required. RDF is now being used to provide such a flexible scheme for cataloguing learning objects on the Web. Indeed, the IMS provides an RDF binding for its Learning Resource Metadata Specification (which is essentially the same as the IEEE LOM Standard).

It is likely that more and more local repositories based on a peer-to-peer (P2P) approach will surface in the future [5]. Individual content producers as well as organizations may opt for this approach since it provides decentralized control over their learning objects. However, one problem likely to emerge with P2P repositories is the proliferation of metadata dialects. This makes the discovery of learning objects particularly difficult. However, approaches such as Edutella are using RDF to allow P2P repositories to interoperate, despite the use of incomplete or different metadata specifications [6].

5. The Need for Shareable Ontologies

The intelligent discovery and assembly of learning objects require information not supported by the current set of elements in the LOM standard. For example, it is necessary for each learning object to specify exactly how that learning object is related to concepts in a particular domain, and the kinds of learning outcomes that are possible in that domain, i.e., an ontology of concepts in a domain. With this kind of knowledge, an agent can compare the course structure developed by a course designer with the learning object based on a common understanding of how they relate to each other. This allows the agent to determine which learning objects are "right" for a particular course. Of course, it is important that the course designer use the same concept ontology in specifying the course structure.

There has been a number of recent efforts aimed at developing ontologies for e-learning, e.g., [7, 8]. There have also been a few recent attempts to link elements in the LOM Standard to specially developed ontologies. This is done through the Meta-Metadata and Classification elements

Ontologies about teaching and learning strategies are also useful since they allow a learning object to specify the kinds of techniques it uses to facilitate learning. Together with concept ontologies, these kinds of ontologies make it possible to personalize instruction to individual learners based on learning preferences, learning designs, etc.

Another kind of ontologies required is for the physical structuring of learning objects [5]. To allow learning objects to be interpreted and rendered consistently in different learning systems, it is important that ontologies be developed for describing the structure of learning objects. This is likely to be different from one discipline to another. For example, the concept of "algorithm" frequently occurs in a computer science learning object, but would be irrelevant in a learning object for a history lesson (unless of course, it's about the history of computing or algorithms).

Developing ontologies is an important aspect of the Semantic Web. However, to be useful, ontologies must be shared so that there is common understanding among learning object producers about what the terms mean. Since it is likely that different groups of people will use different ontologies for learning objects, mappings between these ontologies is also an important requirement.

There are several development efforts currently underway aimed at developing ontologies for the Semantic Web. One of these projects is the DARPA Agent Markup Language (DAML) [www.daml.org], which is a language that can be used for the specification of ontologies. A specific ontology developed for the Web is the Ontology Inference Language (OIL) [9]. OIL uses RDF Schema as a starting point. A European initiative has resulted in the development of an ontology called DAML+OIL, which is a semantic markup language for Web resources based on features of DAML and OIL. Also, the W3C is presently working on its own Web Ontology Language (OWL) that uses DAML+OIL as a starting point. However, there has not been much work so far in the development of ontologies for learning objects, such as described in this section.

The Semantic Web opens up a wide range of possibilities for intelligent discovery and reuse of learning objects. By using shared ontologies, it is possible for software agents to perform most of the processing required in discovering and assembling learning objects. This is a tremendous advantage compared to having humans manually sift through thousands of pages returned by a browser having no semantic understanding of the data, or by a search engine using metadata alone.

Re-visiting the example given in Section 2, consider the instructional designer building a course using learning objects on the Semantic Web. The task is handed over to an agent, which builds a semantic interpretation of the query using an appropriate ontology. The agent goes out into the Web, searching for learning objects that satisfy the query. There may be several learning objects that are linked to the same ontology being used by the agent, so there is a common understanding about what is required. However, some learning objects may actually satisfy the criteria, but use a different ontology. If there are mappings between this ontology and that of the agent, it is possible for some semantic understanding to take place. On the Semantic Web of learning objects, it is thus possible to perform more meaningful searches for learning objects.

6. Comparison with Evolution of Computing

Several years ago, the mainframe computer contained all the processing code and provided services to so-called “dumb” terminals. The mainframe computer also managed the storage and security of data at a central location. For various reasons such as cost, mainframe-based systems eventually gave way to the distributed client/server systems of today. In these systems, processing is distributed to cheap devices scattered all over a network.

The present situation where learning objects are embedded within learning systems is analogous to the situation with mainframe systems in the past. The learning systems of today are essentially centralized, with learning objects (data) managed at a single place by a single system. However, with the growth of learning object repositories on the Semantic Web, this model will give way to a distributed model of processing. The centralized processor in the form of the LCMS must now be able to manage distributed learning objects.

It is interesting to note that in the evolution of computing, distributed processing entered the mainstream after the data became distributed. Taking the parallel further, one can speculate that the next stage of development with learning objects is distributed processing. This will result in the processing that is normally performed within a learning system to be no longer confined to the software in the learning system, but distributed into learning object repositories, and perhaps, into the learning objects themselves.

7. Smarter Learning Objects

On the Semantic Web, learning objects will be much more complex than is currently envisioned today. The physical learning object of today will give way to a more complex conceptual object that contains, or refers to, physical learning resources capable of being rendered in multiple display formats (e.g., HTML, PDF, XML, WML), with links to one or more metadata specifications, and perhaps links to other related learning objects. Moreover, the conceptual learning object will contain links to one or more ontologies that provide sufficient information for reusing the learning object in different contexts.

To keep track of all the different information associated with a learning object, it is important that learning objects become more active, rather than be static chunks of contents as they are in their current manifestations. This leads to the idea of a "smart learning object" [10]. Based on this idea, our vision for smart learning objects is that they should be able to perform many of the tasks typically associated with the LCMS. For example, they should be able to perform intelligent self-analysis when queried by an agent on the Semantic Web. They should also be able to scan the Web looking for related learning objects. This opens up the possibility for more meaningful searches for learning objects by agents, since the learning objects themselves contribute to the search. An active learning object should also be able to interact and learn about new environments (e.g., LCMS) in which they are used, and be able to generate appropriate presentation formats of their content. Thus, their use within an LCMS is not restricted to the capabilities of the LCMS. In essence, much of the functionality of an LCMS will be taken over by the learning objects themselves, resulting in a truly distributed system of active learning objects coordinated by the LCMS. Such a development would parallel the widespread use of distributed computing in the world today.

We are presently working on an implementation of this idea using object-oriented learning objects. In this approach, learning objects are software objects containing data and methods. The data are links to the actual physical resources comprising the learning object, links to metadata, ontologies, conceptualizations of a domain, etc. A learning object may also contain links to other courses or content packages where it has been used before to support searches by software and human agents on the Semantic Web.

The methods in our learning objects are responsible for maintaining and verifying the links to the resources that make up the learning object. They also enable various kinds of queries to be performed based on the data stored.

For example, given a concept map of a domain and a set of pedagogical and other requirements (expressed in RDF), the learning object can determine if it is appropriate for the instructional situation. Rendering methods are also provided to enable the learning object to present itself in different formats such as HTML, XML and PDF.

The object-oriented learning object moves us closer to achieving our vision for an intelligent learning object. Since a learning object is responsible for maintaining all the information available about itself, it can move from repository to repository or learning system and yet keep track of all the information that is vital for its discovery and reuse. Thus, object-oriented learning objects are able to play a more meaningful role on the Semantic Web instead of being viewed as mere static chunks of electronic content.

8. Conclusion

Three enabling technologies for the Semantic Web are XML, RDF and ontologies. Each of these has an important role to play in deploying and reusing learning objects on the Semantic Web. XML is used to markup the structure of a learning object in a machine readable way. It is also used to describe the metadata associated with learning objects. RDF allows the specification of metadata and other information associated with learning objects in a more flexible manner, facilitating the discovery and exchange of learning objects with limited information or more than one metadata specifications. Ontologies allow the specification of concepts in a domain as well as the terms used to markup content in a learning object. Shared ontologies allow for different systems to come to a common understanding of the semantics of a learning object.

Continued research and development effort is needed to facilitate the widespread use of learning objects on the Semantic Web. Of particular importance is the development of ontologies for learning objects. In order for pedagogical agents to find the right learning objects on the Web for a particular instructional situation, appropriate ontologies must be defined for different disciplines in the world. Mappings must also be provided for ontologies defined in different places. Finally, to render learning objects on a display system, it is important for the structural elements of the learning object to be properly understood by a learning system. This too, can be achieved by the development of learning object markup ontologies

References

[1] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web," *Scientific American*, 284 (5), pp. 34-43, May 2001.

[2] T.R. Gruber, "A translation approach to portable ontologies," *Knowledge Acquisition*, vol.5, no. 2, pp.1999-220, 1993.

[3] P. Brusilovsky and J. Vassileva, "Course sequencing techniques for large-scale Web-based education," *Int. Journal of Continuing Engineering Education and Life-long Learning*, to appear.

[4] G. R. Richards, R. McGreal, and N. Freisen, "Learning object repository technologies for telelearning: the evolution of POOL and CanCore," Proceedings of the Informing Science + IT Education Conference, Cork, Ireland, pp. 1333-1341, June 19-21, 2002.

[5] P. Mohan and J. Greer, "Reusable learning objects: current status and future directions", accepted for Ed-Media 2003, Honolulu, Hawaii.

[6] W. Nejdl et al, "EDUTELLA: A P2P networking infrastructure based on RDF," Presented at the World Wide Web 2002 Conference, May 7-11, 2002, Honolulu, Hawaii, USA.

[7] M. A. Sicilia, E. Garcia, P. Diaz, and I. Aedo, "Learning Links: reusable assets with support for vagueness and ontology-based typing", Workshop on Concepts and Ontologies in Web-based Educational Systems, ICCE, December 3-6, 2002.

[8] M.S. Urban and E.G. Barriocanal, "On the integration of IEEE-LOM metadata instances and ontologies," *Learning Technology*, vol. 5, no. 1, 2003.

[9] D. Fensel, F. van Harmelan, I. Horrocks, D.L. McGuinness, and P.F. Patel-Schneider, "OIL: An ontology infrastructure for the Semantic Web", *Intelligent Systems*, pp. 38-44, 2001.

[10] S. Downes, (2002, May) "*Smart Learning Objects*" [online]. Available: http://education.qld.gov.au/staff/learning/courses/sdownesapril.html