Proceedings of the 5th WSEAS International Conference on Telecommunications and Informatics, Istanbul, Turkey, May 27-29, 2006 (pp13-18)

A Framework for Semantic Transcoding of Multimedia Learning Objects

JINAN FIAIDHI and SABAH MOHAMMED
Department of Computer Science, Lakehead University,
955 Oliver Road, Thunder Bay, Ontario P7B 5E1,
Canada
http://flash.lakeheadu.ca/~jfiaidhi
http://flash.lakeheadu.ca/~mohammed

Abstract- The e-Learning technologies makes users capable of accessing a large amount of learning objects created in various parts of the world and across many cultures. Due to the semantic heterogeneity of learning objects built within different cultural contexts, the use of these learning objects by learners is often ineffective. However, developing learning objects in a way that would not only allow their reuse, but promote their repurposing, is a challenging problem. This article proposes an ontology based framework for repurposing multimedia learning object components. Unlike the usual practice where multimedia learning object components are assembled manually, the proposed framework enables on-the-fly access and repurposing of learning object components. The framework supports two processes: the decomposition of learning objects into their components as well as the automatic assembly of these components in real-world applications. For now, the framework supports slide presentations. As an application, we will present in this paper the integration of this functionality in SVG.

Key-Words- Learning Objects, Multimedia Learning Objects, SVG

1. INTRODUCTION

Issues concerning learning object (LO) re-use and repurposing are currently among the most important research topics in the learning technology community [1]. In many cases, we need to repurpose specific parts of a LO instead of the LO in its entirety. Today, repurposing content refers to two broad categories of content recycling. The first is the large-scale transformation of content from one medium, one set of e-learning standards or one application to another. The second category is the small-scale or “one at a time” reuse of individual shared content objects within a course, such as an exercise or an assessment question, which is done on a daily basis as e-learning content is developed and updated. The needs of repurposing multimedia content, in particular, have shaped the technology that is used to create and distribute e-learning. This led to an increasing need for a data integration system that allows transparent and uniform access to spatial data disseminated over a network. The management of multimedia objects, especially digitally represented multimedia objects, can be supported by a vast amount of tools including multimedia presentation software, graphic editors, digital audio or video production systems, audio/video servers or digital broadcast systems. Virtually all existing solutions assume the use of one sort of Multimedia Content Management System [2] which is used to manage large amounts of multimedia objects to facilitate the reuse by content related documentation, automatic content analysis and to support the acquisition, archiving and production workflow for distribution. These systems provide a logically centralized storage and access in multimedia repositories [3]. Accordingly, the Learning Content Management System (LCMS) organize; store and access shared content objects largely to meet these needs. Authoring tools modify, tag, integrate and publish content objects into learning objects in ways that allow them to be used again. Integration and conversion tools, which enable the repurposing from one format to another, such as PowerPoint Presentation to e-learning content, expand the opportunities for faster migration and development. Learning management systems (LMSs) store and host the completed content in cooperation with LCMS, to facilitate the repurposing of content as well. Such centralized content management along with the e-learning standards (e.g. SCORM, AICC and CanCore) have, for the most part, achieved compatibility, and made repurposing easy and commonplace. However, even with this technological harmony, we still face major problems with multimedia integration and delivery. We are particularly in need for a scalable solution that can provide multimedia integration and delivery for ubiquitous computing environments. Such environments heterogeneous and dynamic nature demands a more flexible and effective framework.

The problem of integrating information from multiple Learning Objects sources has received little attention in the multimedia learning communities. As such, the multimedia learning object (MLO) applications must resolve heterogeneities with respect to the schemas and their multimedia data, either to enable their manipulation or to enable the translation of multimedia and queries across the schemas. At present, there are various multimedia frameworks, such as those used for video conferencing (H.323, SIP, Access Grid http://www.sipcenter.com/sip.nsf/html/SIP+and+H.323, http://www.accessgrid.org), and for multimedia courseware presentations (e.g. PowerPoint, Acrobat, Director, Corel, Lotus, Harvard). All such frameworks can not usually interact with each other or share data among each other. Traditionally this problem has been partially solved using complicated multimedia integration systems that provide a uniform interface to the database sources as well as on developing dedicated multimedia transcoders using techniques such as Bitrate Reduction or Multi-Channel Output Separator. Although there are many trials on building multimedia data integration systems that provide a uniform query interface to the different multimedia sources (e.g. Global MMCS [4], JETS[5]) and on dedicated transcoders (e.g MPEG2 to MPEG4 Transcoder [6]), these trials raises a new design issues for application developers. Questions on media composition, media translation, media synchronization, data formats, and user interfaces must be examined in the light of the capabilities of multimedia platforms. The need for a smart, flexible system for automatic extraction, analysis, integration and distribution of multimedia content is self evident. Such multimedia manageability encompasses many other issues such as scalability. Scalability is an important aspect of multimedia management so as to allow MLOs to support other user needs and purposes. This criteria means that with scalability we need to support semantic transcoding of multimedia while preserving their consistency, both in terms of information and its presentation. Generally, a semantic transcoding management policy is composed of multiple distributed components which place great weight on the multimedia identification and integration procedures. With such policy, Multimedia design can be reduced to the process of choosing a presentation form which can be mapped to a set of domain concepts which you wish to communicate to users so that they can use the concepts to perform a task as effectively and efficiently as possible in their own domain. However, one of the major choices in multimedia design is to choose how much of the design process takes place off-line by a skilled human designer, and how much is performed automatically by the system. The consequences of this choice will determine the role of the designer and the concomitant interactions with the constraints on multimedia design within the required e-learning environment. Most researchers who have a vision for a semantic transcoding policy uses an open source SMIL framework for constructing flexible MLOs [7,8,9,10]. SMIL supports four constructs: layout, timing, hyperlinking and tailorability of the presentation, while the human designer chooses the content of a presentation. At the end the SMIL presentation designer holds all knowledge of the task and domain, using it to describe the presentation using the four constructs provided by the language. The SMIL presentation is sensitive to available bandwidth, presentation station capabilities and user attributes which can be used at run-time to select between alternatives specified by the designer, but otherwise all decisions are made by the designer at authoring time.

Current SMIL-Based systems provide databases and repositories to index MLO collections with user-defined keyword annotations and media-specific metadata import. Although a step in the right direction, these systems use proprietary databases without access to the stored organisation structures or to metadata outside the designed application. In this context the Semantic Web provides data models to describe Web resources (e.g.,MLO) with application-independent languages such as the Resource Description Framework (RDF) [11] which are also applicable to multimedia resources. The vocabulary used to describe documents can be specified in terms of ontologies, where each description term and its semantic relation to other terms are defined. " Ontologies provide a shared and common understanding of a domain that can be communicated between people and application systems" [12], and thus ontologies facilitate the sharing and exchange of information about multimedia between applications and users. Accordingly, a multimedia ontology comprises a shared vocabulary to describe multimedia documents and their organization in a structured way such that users and applications can process the descriptions with reference to a common understanding specified in ontologies.

2. Multimedia Ontology

At the present, users typically arrange their multimedia collections into file systems/databases which provide poor naming mechanisms and hierarchical directory structures for organization, composition and searching. In particular, this approach has the following drawbacks:

• The categorization depends on the used classification hierarchies;

• The logical organization strictly depends on the physical storage system;

• Identification based on file names alone is often not globally consistent (e.g., duplicates are possible);

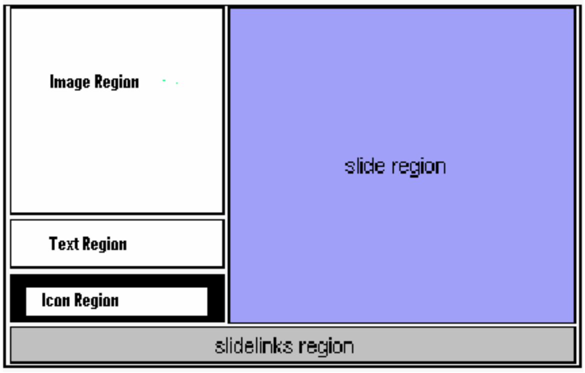
• The semantic content of multimedia objects is difficult to represent and manage.

Semantics of multimedia materials are very hard to capture either by manual or automatic way: these semantics may be viewed as the set of terms created or linked in the practice, which forms the multimedia ontology of the discourse. A complete and extensible ontology that expresses the basic concepts that are common across a variety of domains and media types and that can provide the basis for specialization into domain-specific concepts and vocabularies, is essential for well-defined mappings between domain-specific knowledge representations (i.e., metadata vocabularies) and the subsequent building of a variety of services such as cross-domain searching, tracking, browsing, data mining and knowledge acquisition. As more and more communities develop metadata *application profiles* which combine terms from multiple vocabularies (e.g., Dublin Core, MPEG-7, MPEG-21, CIDOC/CRM, FGDC, IMS) such a core ontology will provide a common understanding of the basic entities and relationships which is essential for semantic interoperability and multimedia transcoding. Typically, ontology consists of *concepts*, concept *properties*, and *relationships* between concepts. In a typical ontology concepts are represented by terms. In a multimedia ontology concepts might be represented by multimedia entities (images, graphics, video, audio, segments, etc.) or terms. Such core ontologies have also concentrated on defining XML representations of their description schemes using the XML Schema languages (e.g, RDF Schema [11], DAML+OIL [12], Web Ontology Language (OWL) [13]). The knowledge representation provided by such ontologies can be used to develop sophisticated services and tools which perform knowledge-based reasoning, knowledge adaptation, knowledge integration and sharing, transcoding and knowledge acquisition, specifically for semantically-rich multimedia content.

To this end, several approaches in literature which address the problem of building multimedia ontologies to enable the inclusion and exchange of multimedia content through a common understanding of the multimedia content description and semantic information. As a well-known fact, the *MPEG-7* (http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm) and *TV-Anytime* (http://www.tv-anytime.org/) are the most notable standards that constitutes the greatest effort for multimedia description. MPEG-7 provides in the Semantic Part of the MPEG-7 MDS, the complex data types needed for the complete semantic description of audiovisual content as a set of *Description Scheme (DSs)*, while TV-Anytime provides only keyword-based semantic description capabilities. Clearly MPEG-7 is far more complicated than TC-Anytime [14]. Actually MPEG-7 is implemented by XML Schemas. The set of MPEG-7 XML Schemas defines 1182 elements, 417 attributes and 377 complex types. The size of this standard makes it quite difficult to manage. However, these standards do not propose a concrete methodology and language for the integration of domain specific knowledge for the multimedia content as they provide only general-purpose structures for metadata representation. Moreover, domain-specific knowledge for a specific domain related to the content of a MLO presentation may be described in a well-accepted ontology description language like OWL [15] which is independent of MPEG-7 and TV-Anytime.

The present paper described a framework for transcoding of MLOs stored in a meaningful SVG based on finite number of key terms and patterns frequently met in the domain of MLO presentation. This constraint allows us to minimize the number of key terms to be recognized in MLOs. The advantage of having a meaningful XML representation in the form of SVG is in its easy conversion to other formats, including those suited for building ontologies (such as OWL, RDF, RuleML, etc.). Also the meaningful XML documents can easily be integrated with some existing XML-based repositories of learning objects (such as SCORM, IEEE LOM, CanLOM, etc.) Finally, the meaningful XML already contains the multimedia learning object's metadata (context), which allows applying the approach of context mediation to attain semantic interoperability between semantically heterogeneous MLOs and facilitating meaningful delivery of MLOs to the learner's contexts.

In our proposed framework, the process of building and integrating ontologies of the MLOs consists of two main processes: the *composition of MLO* (i.e. the integration step) and the *decomposition of MLOs* into their components (i.e. the extraction steps). However, a *preprocessing* step might be added as an initial step required for creating new MLOs. The new MLOs can be created with existing, new or modified ontology. The most important factor is that these ontologies must have a uniform hierarchical structure (e.g. tree-like). Indeed, most of the tools that utilize ontologies, such as ReTAX, Protégé, OntoEdit, OilEd, WebODE and OntoRAMA, offer a facility of hierarchy viewing to support the user to build and edit ontologies. A hierarchical view of ontology seems to be a good way to give the user a quick overview of the selected ontology. Figure 1 illustrates a general view to the hierarchy of a typical MLO presentation.



(a) A Graphical Sketch for the MLO Hierarchy

```
<multimedia presentation>
<head>
 <layout>
        <root-layout width=851 height=95>
        <region id=”imageregion” top=5 width=350 height=262>
        <region id=”slideregion” top=5 left=352 width=499 height=355>
        <region id=”textregion” top=267 width=350 height=45>
        <region id=”iconregion” top=313 width=350 height=45>
        <region id=”slidelinksregion” top=360 width=851 height=35>
 </layout>
</head>
<body>
 <par>
  <img src= “http://... “ region=”imageregion”>
        <img src= “http://... “ region=”slideregion”>
        <textstream src= “http://... “ region=”textregion”>
        <img src= “http://... “ region=”slidelinksregion”>
        <textstream src= “http://... “ region=”imageregion”>
</par>
</body>
</multimedia presentation>
```

(b) XML Representation of the MLO hierarchy.

Fig. 1 MLO Presentation Hierarchy.

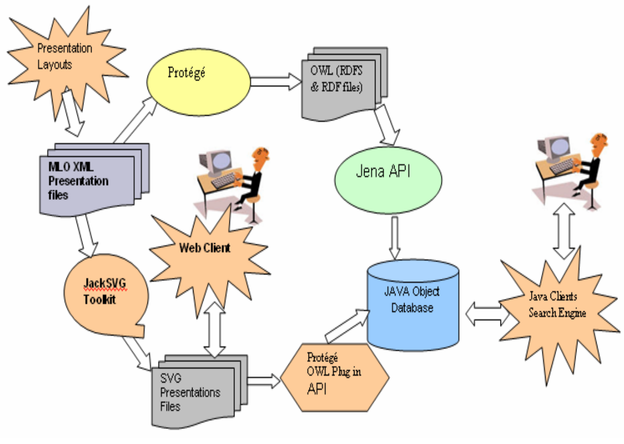
Actually after evaluating a number of different ontology representation tools we chose Protégé (http://protege.stanford.edu/) for two reasons:

1. The user interface provides an environment that can easily be learned and remembered by the domain experts.
2. The ever-growing number of plug-ins provides Protégé with considerable versatility.

Protégé is based on Java, is extensible, and provides a foundation for customized knowledge-base that can help in extracting the ontology structure. Protégé supports Frames, XML Schema, RDF(S) and OWL. It provides a plug-and-play environment that makes it a flexible base for rapid prototyping and application development. The second reason proved especially valuable in migrating the data content from the XML hierarchical structure to a format readable and editable by Protégé. Especially with Protégé-OWL plugs in, we can have services for ontology model representation, parsing, database persistence, querying and some visualization tools. Protege-OWL always had also a close relationship with Jena (http://jena.sourceforge.net). Jena provides rule-based reasoners are able to provide semantic entailments for ontologies using OWL. Moreover, Jena provides a query engine.

However, composing the multimedia content from the MLO presentation hierarchy can be made by creating an SVG file through writing a Java code that makes calls to the database, gets data, and writes an SVG file (e.g. based on tools like Apatche Batik API). In our framework, on the other hand, we eliminates Java coding at the composition stage; instead, we use JackSVG toolkit (titanium.dstc.edu.au/xml/jacksvg/) as a middle tier that externalizes both database calls and transformation of data from the multimedia database into SVG. The reason on transferring the MLO into an SVG file is a straightforward: SVG provides a powerful and rich language to describe presentational content and it is a platform independent presentation media. SVG is the new W3C open standard. Moreover, the use of SVG leads to clear separation of data-access logic from presentation logic. SVG allows client interaction with the MLO as well as it opens the doors for cross client communication through network sockets. Thus, allowing the SVG applications to share data in real time with other clients or applications, enabling the user besides viewing the presentation also to interact with it in a multi user environment. Actually, SVG is based on the DOM model and hence it can be transcoded and rendered at any viewing platform. So we can say with great confidence that SVG is more than a simple collection of slides which are only one part of the whole presentation. Other parts of SVG are the ability for the viewer to collaborate with others and contribute content to the presentation. SVG replaces server-side multimedia file creation or applet-based graphics with client-based rendering of multimedia, and is likely to revolutionize the way Web multimedia are rendered, stored, manipulated, and associated with content [16]. Figure 2 illustrates our framework in relation with the notion of multimedia ontology.

Fig. 2 The MLO Semantic Transcoding Framework.



3. MLOs Semantic Transcoding

In this section, we are shading some light on the problem of semantic transcoding of MLOs given a framework like the one we proposed in Figure 2. Indeed, semantic heterogeneity is a key problem in any data sharing system, be it a federated web services, data integration system, message passing system, peer-data management system, or MLOs translation and transcoding. At an abstract level semantic transcoding identify concepts in models that semantically correspond to each other. By doing so, they bind together data available in different ontology models, enabling migration, reasoning and query processing across different ontologies. The data sources involved are typically designed independently, and hence use different schemas. To obtain meaningful inter-operation, one needs a semantic mapping and transcoding between the different schemas or ontologies, (i.e., a set of expressions that specify how the data in one multimedia corresponds to the data in the other.)

Schema matching is inherently a difficult task to automate mostly because the exact semantics of the data are only completely understood by the designers of the schema, and not fully captured by the schema itself. In part, this is due to the limited expressive-power of the data model, and often is further hindered by poor database design and documentation. As a result, the process of producing semantic mappings requires a human in the loop and is typically labor-intensive, causing a significant bottleneck in building and maintaining data sharing applications. Schema matching (a.k.a ontology alignment) has received steady attention in the database and AI communities over the years [17]. A key conclusion from this body of research is that an effective schema matching tool requires a principled combination of several *base techniques*, such as linguistic matching of names of schema elements, detecting overlap in the choice of data types and representation of data values, considering patterns in relationships between elements, and using domain knowledge. However, current solutions are often very brittle. In part, this is because they only exploit evidence that is present in the two schemas being matched. These schemas often lack sufficient evidence to be able to discover matches.

As would be expected, people have tried building semiautomated schema matching systems by employing a variety of heuristics as well as syntactical and semantically mappings. However, these techniques are not based on the different ontologies structures. In this direction we find that the use of the "Schematron toolkit" (http://www.schematron.com/) will enable us to achieve this goal. This toolkit differs in basic concept from other schema matching techniques in that it not based on grammars but on finding tree patterns in the parsed document and use it as a base for comparison. This approach allows many kinds of structures to be represented which are inconvenient and difficult in grammar-based schema languages. Schematron is built on XPath and XSLT expression languages. The Schematron allows you to develop and mix two kinds of schemas:

- *Report* elements allow you to diagnose which variant of a language you are dealing with.
- *Assert* elements allow you to confirm that the document conforms to a particular schema.

The schema matching primitives in the Schematron are based on two based simple actions:

- First, **find** a context nodes in the document (typically an element) based on XPath path criteria;
- Then, **check** to see if some other XPath expressions are true, for each of those nodes.

With tools like Schematron we can perform ontology translation by using techniques like ontology merging and automated reasoning [18,19]. Our focus need to be on inferences from facts expressed in one ontology to facts expressed in another. The *merge* of two related ontologies is obtained by taking the union of the terms and the axioms defining them, using SVG namespaces to avoid name clashes. Once the merged ontology is obtained, ontology translation can proceed without human intervention.

4. Conclusions

This article introduced an overview of a framework for MLO presentations transcoding. The key elements of this framework are based on techniques to enhance MLO content using ontologies. In the previous work on using ontologies to describe MLOs, researchers have built ontologies for description of metadata. However, these ontologies do not improve an LO's content. We suggest creating MLOs that have content marked up in accordance with domain ontologies. Accordingly, MLOs can be used not only as learning materials, but they can also be used in real-world applications. For this purpose, the framework suggest the use of Protégé toolkit (for MLO ontology construction), the use of Protégé-OWL API and Jena API to dynamically represent the ontologies as related Java classes (for MLO representation and retrieval), the use of JackSVG to convert the primitive XML description of a MLO presentation into a dynamic SVG presentation (for easy multimedia content rendering), and finally the use of Schematron along with *ontology translation* techniques (for MLO transcoding and translation among the various styles of multimedia presentations.) We are currently experimenting with various scenarios for implementing the proposed framework.

References

[1] K. Verbert, J. Jovanovic, D. Gaševic, E. Duval, Repurposing Learning Object Components, OTM 2005 Workshop on Ontologies, Semantics and E-Learning, November 3 2005, Agia Napa, Cyprus

[2] T. Kleinberger and P. Müller, Content Management for Web Based Education. Webnet 2000 Conference on the WWW and Internet, San Antonio, USA, 30.10.2000 - 4.11.2000, 329-334.

[3] T. Kleinberger, L. Schrepfer, et al., A Multimedia Repository for Online Educational Content. ED-MEDIA 2001 World Conference on Educational Multimedia, Hypermedia and Telecommunications, Tampere, Finnland, June 25-30, 2001, 975-980.

[4] Geoffrey Fox, Wenjun Wu, Ahmet Uyar, Hasan Bulut, Shrideep Pallickara**,** Global Multimedia Collaboration System, 1st International Workshop on Middleware for Grid Computing Co-located with Middleware 2003, Rio de Janeiro, Brazil , June 15-17 2003

[5] S. Shirmohammadi and N.D. Geroganas, "JETS: a Java-Enabled Telecollaboration System", proc. IEEE International Conference on Multimedia Computing and Systems, ICMCS'97, Ottawa, Canada, pp. 541-547, June 1997.

[6] Kalva, H.; Vetro, A.; Sun, H., "Performance Optimization of the MPEG-2 to MPEG-4 Video Transcoder", SPIE Conference on VLSI Circuits and Systems, Vol. 5117, pp. 341-350, May 2003 (SPIE Proceedings, TR2003-057)

[7] Jane Hunter and Suzanne Little, Building and Indexing a Distributed Multimedia Presentation Archive Using SMIL. 415-428 5th European Conference, ECDL 2001, Darmstadt, Germany, September 4-9, 2001

[8] Dick Bulterman, Lynda Hardman, Sjoerd Mullender, Lloyd Rutledge, Martin Prime, Michael Wilson GRiNS: A GRaphical INterface for Creating and Playing SMIL Documents. *Proceedings of the EMMSEC 98 conference* Elsevier/CEC, Brussels, (August 1998)

[9] C A Macnee, et. al, Presenting Dynamically Expandable Hypermedia, *Information and Software Technology Vol.* 37, p.339-350 Elsevier, (1995)

[10]. J-L Binot, P Falzon, R Perez, B Peroche, N Sheehy, J Rouault, M D Wilson Architecture of a multimodal dialogue interface for knowledge-based systems *Proceedings of Esprit 1990 Conference* p.412-433 Kluwer Academic Publishers, Dordrecht, (1990)

[11] F. Manola and E. Miller, RDF Primer. W3C Recommendation Series Editor: Brian McBride (Hewlett-Packard Laboratories). 10-February-2004. http://www.w3.org/TR/rdf-primer/

[12] M. Klein, D. Fensel, F. van Harmelen, and I. Horrocks: The relation between ontologies and schema-languages: Translating OIL-specifications in XML-Schema. In Proceedings of the Workshop on Applications of Ontologies and Problem-solving Methods, 14th European Conference on Artificial Intelligence ECAI'00, Berlin, Germany August 20-25, 2000.

[13] M. K. Smith, C. Welty, and D. L. McGuinness, "OWL Web Ontology Language Guide," 2003. http://www.w3.org/TR/2003/PR-owl-guide-20031215.

[14] Anastasia Karanastasi , G. Kazasis and Stavros Christodoulakis, A natural language model for managing TV-Anytime information in mobile environments, Personal and Ubiquitous Computing**,** Volume 9 , Issue 5 (September 2005), pp262 – 272, 2005

[15] Tsinaraki C., Polydoros P., Christodoulakis S., *"Integration of OWL ontologies in MPEG-7 and TVAnytime compliant Semantic Indexing"*, In the proceedings of the 16th International Conference on Advanced Information Systems Engineering (CAiSE), Riga, Latvia, June 2004

[16] Dragan Gasevic, Jelena Jovanovic, and Vladan Devedzic, Enhancing Learning Object Content on the Semantic Web, Fourth IEEE International Conference on Advanced Learning Technologies (ICALT'04) pp. 714-716

[17] Erhard Rahm, Philip A. Bernstein, A survey of approaches to automatic schema matching, The VLDB Journal 10: 334–350 (2001).

[18] Dejing Dou, Drew McDermott, and Peishen Qi,, Ontology Translation on the Semantic Web**,** In R. Meersman et al. (Eds.): CoopIS/DOA/ODBASE 2003, LNCS 2888, pp. 952–969, 2003.

[19] Dejing Dou, Drew McDermott, and Peishen Qi**,** Ontology translation by ontology merging and automated reasoning. In *Proc. EKAW2002 Workshop on Ontologies for Multi-AgentSystems.* pp. 3-18.