Maritime Transportation and the Semantic Web

Raphael Malyankar
Dept. of Computer Science and Engineering
Arizona State University, Tempe, AZ 85287-5406, USA.
E-mail: rmm@acm.org

1 Introduction

My work involves the creation of a computational ontology for maritime information, and the use of this ontology in markup and integrated retrieval from sources of maritime information. Ontological information is acquired from multiple sources, including standards documents, database schemas, lexicons, collections of symbology definitions, and also from semi-structured documents. The computational ontology thus created is being used to create an XML-based markup language (Maritime Information Markup Language — MIML) for tagging information within this domain. This statement describes progress on this project and my views on its relation to the Semantic Web.

The ultimate purpose of this project is to help upgrade maritime information distribution for the Semantic Web. A number of disparate distribution channels are currently being used, ranging from paper documents published once every few years, to radio distribution of warnings and weather forecasts nearly in real-time. Early, accurate, and integrated information is very important to the waterborne transport industry for reasons of safety as well as cost, and there are a number of ongoing efforts concerning upgrading information distribution, but distributing and combining information from different sources is still difficult and resource-intensive. Semantic Web technology should help with these problems, while still allowing leverage of existing distribution channels. The intention is to get information into end-users' hands in a *usable* form, to make it easy for producers of information to 'maintain' (update, revise, or extend) the information they produce, and to make it possible for application developers to write applications that can process it.

The work is part of a larger interest in the representation and distribution of geospatial information, particularly that relating to navigation for all modes of transport. It is related to my view of practical uses for the Semantic Web; I foresee the most important practical use of the Semantic Web in the near future to be the exchange of 'knowledge-at-a-distance' for transportation and logistics, because a large part of this information is not available or cannot be confirmed until shortly before it is needed, and further it may not be in exactly the form desired by the user; but when it does become available, it needs to be integrated into existing plans and local knowledge bases very quickly.

2 Ontology and Markup Language

Ontological knowledge was extracted from various sources, ranging from standards created by standards bodies to semi-structured 'content' documents. Extraction was done by a mixture of semi-automated and human means, and is described elsewhere[1]. The ontology learning stage resulted in a collection of overlapping sub-ontologies, which were merged into one large taxonomy. This taxonomy is currently being used for two purposes: first, as a kind of hierarchical index into a multifarious, distributed, knowledge base consisting of knowledge sources of different kinds — digital charts, web sites, programs, and marked-up text documents; and second, as a basis for defining a markup language for our domain. Reasoning and inference will be investigated soon.

In its 'hierarchical index' role, it is being used to map concepts from the 'user domain' to the 'knowledge source domain', i.e., map concepts to (possibly multiple) information sources. A prototype information

[1] URL: www.isi.edu/dgrc/dgo2001/papers/session-1/malyankar.pdf

```
<Chart>
 <ChartNumber>18773</ChartNumber> <ChartNumber>18772</ChartNumber>
 <Description>San Diego Bay is where California's maritime history...</Description>
 ...
  <PierArea> ...
   <Pier  name ="Tenth Avenue Marine Terminal">
    <Berth  name="Berths 1 and 2">
     Concrete bulkhead, 1,170 feet of berthing space; 27 feet alongside...
    </Berth>
    <Berth  name="Berths 3 and 6"> ... </Berth>
    <Berth  name="Berths 7 and 8"> ... </Berth>
    ...
</Chart>
```

Figure 1: Marked-up fragment of Chapter 4, Volume 7 of the *Coast Pilot*

retrieval application has been constructed; it retrieves information from different sources in response to a user query; for example, a request for weather information gets weather forecasts from Web sites as well as notes from the *Coast Pilot* (a text document containing information of interest to mariners sailing along the U. S. coastline). URIs are used to link knowledge sources to ontological concepts.

The Maritime Information Markup Language (MIML) is an XML-based language used for text markup in the prototype described. Figure 1 contains an example of its use. Some MIML tags are derived (via the ontology) from direct sources, such as an IHO (International Hydrographic Organization) standard, a lexicon published by NOAA (National Oceanic and Atmospheric Agency), feature classes in a sample Digital Navigation Chart, etc. A second class of tags was also needed to denote information elements within the document, for example the <Chart> tags, which denote sections that pertain to a specific (identified) nautical chart in the NOAA chart numbering system, and the <Description> tag, which is used to denote general text information that cannot be placed into a more specific category. These were invented as necessary. The DTDs (Document Type Definitions) for the Coast Pilot were prepared 'by hand'.

3 Technological and Research Issues

Some general issues important to my work are the use of Semantic Web technologies for distributing and processing location-dependent information, reasoning about such spatially dependent information, and integrating this information into local knowledge bases, especially for information that needs to be distributed in near-real-time, such as weather warnings. The use of spatial reasoning for navigation of all types also needs exploration, especially reasoning for navigation in a dynamic domain where navigation planning is affected by dynamic processes as well as mobile entities and obstacles, and 'feature-based' reasoning, where a navigator must use landmarks and required routes. Further issues are the retrieval and indexing of all kinds of information, ranging from text to pictorial and diagrammatic representations. Last but not least, it is extremely important to create common standards for information markup, and gain the support of governing bodies or organizations; without these, widespread acceptance of Semantic Web technologies will not be possible in these application domains.

Acknowledgements

This work was partially supported by the National Science Foundation under grant EIA-9983267, NOAA, Maptech, and the U.S. Coast Guard. The opinions, findings, conclusions and recommendations expressed in this material are those of the author and do not necessarily reflect the views of these entities.