Information Platform for Scientific Knowledge Sharing

Toshihiro Ashino[1], Takayuki Saito[2]

[1] IMS Lab. Inc., [2] IMSLAB

http://www.interwall.org/~ashino/

Because of numerical experiment has became a common and important method for science and technology today, a common platform for archiving, organizing and retrieving scientific knowledge is required. XML tag set for managing computational science knowledge and prototype implementation of platform will be presented.

Introduction

In computational science and engineering research, input data and output data have been regarded as temporary data and generated for each transaction. But these days, numerical experiment has become a common and important method for science and technology, sharing programs and related knowledge is becoming important. XML provides flexible framework to represent knowledge about computational science and engineering, include program itself, documents, input and output data sets and their relation. In this paper, prototype platform for scientific knowledge management is discussed.

Sharing Knowledge with XML

XML is regarded as the standard to exchange data over the Internet. With XML, we can define a vocabulary for specific field, this means, the meaning of document can be processed by program when annotated with appropriate XML tag set. Also in the computational science and engineering field, common data format can be defined with XML, and related documents, programs, images etc., can be linked each other by hyperlink. "netCDF" is common data exchange format for binary data within atmospheric research community, but netCDF provides no standard for parsing tags. There are so many softwares for parsing and editing XML and with related standards can be employeed, for example, equation can be embedded on the document with MathML.

Prototype System

A consortium had been established Nov. 2000, for discussing XML tag set standardization for computational science. Consortium - name in English is not yet fixed - includes researchers from several fields, computational fluid dynamics, life science, etc.

A prototype system of a platform for scientific knoledge sharing had been developed and tested on FY 1999. This system uses Object-Oriented database management system in order to manage XML documents and each document linked by hyperlinks. XML documents are divided into two groups, one describes single case of calculation, input and output data, related document, etc., the other group describes program itself, program source code, manual and sample input data.

Prototype system defines many XML tags and these are classified as follows,

- System, Database Management tags
  * bibliographic data
  * document structure
  * communication protocol
  * tag dictionary
- Common attribute tags of data
  * physical constants
  * physical properties
  * data types
  * unit system
- Field specific tags
  * biopolymers
  * computational fluid dynamics
  * unstructured mesh generation
  * computational materials science

Prototype system had been tested few months. Distributed servers had been sited on several organizations and frontend softwares written with Java had been distributed. With this system, knowledge about numerical experiment had been shared over several research groups, especially settings of empirical parameters had been refered frequently and activated collaboration.

Conclusion

Sharing knowledge of computational science and engineering is important to enhance collaboration of researchers. Consortium members are planning to organize special interest group to discuss knowledge exchange for some fields, and to define common problems, i.e. common tag definition problem like unit system description, technical problems, like numeric data handling in XML, effective XML document management, etc.