Safety Markup Language: Concept and Application

Chin-Feng Fan¹ and Swu Yih²

¹ Department of Computer Engineering and Science
Yuan-Ze University, 135 Far East Road, Chung-Li, Taiwan
csfanc@saturn.yzu.edu.tw

² I&C Department, Institute of Nuclear Energy Research
P.O. Box 3-11, Lung-Tang, Taiwan

Abstract. This paper proposes a method to expedite hypertext construction and improve document reading efficiency by using meaningful tags. Hazard Life Cycle is proposed as the common semantic framework for documents in safety engineering practice. Under this framework, we developed Safety Markup Language (SML) to annotate the major concepts in software-related nuclear regulation. A computer tool has been constructed to convert these SML tags into desired hyperlinks for review purpose. This approach reduces the manual effort in hyperlink construction, and supports information retrieval in a concept unit, which is closer to human cognition than that obtained from a conventional approach. Potential improvements achieved by this SML-based method include efficient checking of information completeness, tracing of review issues, and reduction of clerical work in license review.

1. Introduction

Document reading has become a more and more important activity for safety engineering practitioners in such tasks as licensing review, accident investigation, and product liability argument. Safety engineers can handle documents of simple systems effectively, but for complex systems like nuclear power plants, the huge amount of inter-related documents has become an enormous burden for those who have to read and process them. In certain cases, the overwhelming documents have created an unbalanced situation between document load and safety engineers’ document processing capability. Such an unbalanced situation not only deteriorates safety engineers’ performance but may also jeopardize safety of the target equipment. Effective document handling tools and techniques are urgently needed.

Computer supported hypertext is promising in enhancing people’s document reading capability. It provides a structure better matching human reading and cognition process than conventional paper-based documents. However, current hypertext approach is mostly address-based, and the constructing of address-based hyperlinks is very tedious and time-consuming. This paper proposes an efficient way to expedite hypertext construction by using meaningful tags. The rationale is that

¹ This work was supported in part by National Science Council, Republic of China, under the grant no. NSC87-2218-E-155-014-NU and NSC88-2213-E-155-002.
since the contents of each document is composed of inter-related domain concepts, these concepts can be organized into a semantic framework. This framework is shared by all concerned documents. Therefore, we can assign names/tags to these domain concepts, and then hyperlinks can be established among document paragraphs that have the same or related tags. Furthermore, the constructed hypertext system can be retrieved in a concept unit; thus, the presentation mechanism is closer to human cognition mechanism than that of a conventional approach. A computer tool assisting to realize this approach will be needed to help safety engineers to cope with complicated documents. The unbalancing problem between overloaded documents and people’s reading capability will then be alleviated.

In this paper, a Safety Markup Language (SML) is defined to represent the elements and dependency relations of the common semantic framework of software-related nuclear regulation. Application-specific processing of the SML tags will then be applied to form an efficient hyper-document. In the following, we will first discuss related background, and then present our approach, followed by the case study in software-related nuclear regulation. Improvements achieved by this SML-based approach are also discussed.

2. Background

This research uses markup to construct hyperlink and provides efficient reading. Thus, we will discuss hypertext and markup in this section. Hypertext refers to electronic documents in which information is arranged and accessed in a nonlinear fashion [1,3,6]. The interconnection between chunks of documents is usually defined by pre-defined, address-based hyperlinks. However, the creation of address-based hyper-links among various documents is usually tedious and needs to be done manually.

Markup traditionally refers to the annotation or instructions for formatting or typesetting. The names for the annotation are called tags. Such formatting languages include Unix nroff and troff. Later on markup languages become more general; besides formatting, they can provide annotation for text syntactical structures such as title and paragraph. SGML[4], XML[2] are such general markup languages. HTML, providing formatting markup, is a subset of SGML and popular in constructing homepage on WWW. There are a few famous domain specific markup languages, such as Chemical Markup Language (CML)[7] and Mathematical Markup Language (MathML)[5].

3. Concept-Based Approach

To enhance reading efficiency, we have investigated human reading and identified key factors affecting reading efficiency. The top portion of Fig. 1 shows the typical reading process. The underlining process is the concept transfer process from the writer to the reader. If the writer’s concepts can be represented explicitly and flexibly, then reading efficiency can be enhanced. The key tasks are how to extract, present, and then manage the concepts as well as their dependencies. Our approach to these key tasks is shown in the lower part of Fig. 1. To extract the concepts to transfer, we