A Generic Model for Universal Data Storage and Conversion and Its Web Based Prototypical Implementation

Andreas Unterweger, Salzburg University of Applied Sciences, Austria
Bernadette Himmelbauer, Salzburg University of Applied Sciences, Austria
Simon Kranzer, Salzburg University of Applied Sciences, Austria
Peter Ott, Salzburg University of Applied Sciences, Austria
Robert Merz, Salzburg University of Applied Sciences, Austria
Gerhard Jöchtl, Salzburg University of Applied Sciences, Austria

ABSTRACT

This paper presents a prototypical, Web based data conversion framework and its underlying data representation principles which allow conversions from and to any data format. Therefore, a data model is proposed which allows storing values of arbitrary types, including inter-data dependencies and meta information. Furthermore, an Extensible Markup Language (XML) based model to describe data formats is provided which allows specifying programs to convert data represented in existing formats both from and to the proposed data model. It will be shown that these programs are Turing complete, thus allowing the same arbitrarily complex conversions which are possible with Extensible Stylesheet Language Transformations (XSLT) or the C programming language. Finally, the components of a prototypical Web based implementation in form of a validator, a data converter and a data generator are described. In combination with a data editor, parts of this prototypical implementation are already employed in several use cases in the industry and other research projects to transform data between different formats.

Keywords: Continuous Data Integration, Data Conversion, Data Model, Data Storage, Data Transformation, Graphical User Interface, Turing Completeness, Web Based, XML Model

INTRODUCTION

Whenever huge quantities of information from industrial applications need to be stored, transformation and manipulation of data is a complex issue. Vast amount of data combined with proprietary data formats, which are generated from different data sources, pose a great challenge for data handling. To address this issue, we present a fully integrated solution.
which enables storage and transformation of arbitrary data formats.

Our approach is based on two essential models: a generic data model and an XML model. The generic data model allows storage of arbitrary data formats, including meta information and interdependencies. The XML model specifies data formats and serves as transformation language from and to the original data representations. In addition, a universal data converter and a generator are implemented in a prototype framework. Figure 1 shows our complete conversion approach. Based on the XML model, a program for transformation only needs to be specified once for a particular data format. The same model will be used for both conversion and generation.

Similar to our XML model, a transformation language has been realized in XSLT by the W3C (2004) and allows the conversion of arbitrarily complex data formats, the XSLT version requires a different data format specification for each transformation direction (input and output). Furthermore, the XSLT implementation does not specify how storage inside the database is handled, while in our approach storage structures are explicitly defined by the generic data model. biXid, a bidirectional transformation language, is shown by Kawanaka and Hosoya (2006) and designed for the transformation of XML formats only, as is DSLTrans proposed by Barroca, Lucio, Amaral, Felix, and Sousa (2010), which however, is Turing incomplete.

In addition to the lack of a generic data model (see XSLT above), parser generators like ANTLR (Parr, 2007) cannot be used for text generation, while text processing software like awk (IEEE, 2004) or commercial data converters like Altova MapForce (Altova, 2011) impose limits on data formats and/or the complexity of the conversion. As this makes them and all similar solutions either non-universal or not as flexible as our approach, they are not reviewed in detail.

The paper is structured as follows: first, the XML model as the core part of our approach which allows for data format description and conversion is explained. After showing the Turing completeness of said XML model in the subsequent section, the generic data model for storing data values, their dependencies and the corresponding meta information is described. The Web based prototypical implementation of a universal data converter and data generator based on the two models, i.e. the data and the XML model, is described in the “Implementation” section.

THE XML MODEL

In order to specify the format of data to be parsed or generated, we developed an XML model, implemented in form of an XML schema (W3C, 2004) whose complexity is sufficient to model...
Related Content

A Unified Approach to Uncertainty-Aware Ubiquitous Localisation of Mobile Users
www.igi-global.com/article/unified-approach-uncertainty-aware-ubiquitous/65067?camid=4v1a

A Prediction-Based Flexible Channel Assignment in Wireless Networks Using Road Topology Information
www.igi-global.com/article/prediction-based-flexible-channel-assignment/2617?camid=4v1a

On the Prospects and Concerns of Pattern-Oriented Web Engineering
www.igi-global.com/chapter/prospects-concerns-pattern-oriented-web/40423?camid=4v1a
Librarian without Building in an E-Learning Environment: Needed Skills, Challenges, and Solutions