

XML and Web Technologies

INFO-H509

XML Schema Definition

16 April 2021

BAKKALI Yahya : 000445166

HAUWAERT Maxime : 000461714

UNIVERSITÉ LIBRE DE BRUXELLES (ULB)

Contents

1 Introduction 2

2 XML Schema 2

2.1 XSD Version 2

2.2 ISBN 2

2.3 Currency 2

2.4 Genre 2

2.5 Price 2

2.6 Volume 3

3 Miscellaneous 3

4 XML Examples 3

A Relations 4

B Tree 5

1 Introduction

A bookshop has two different departments a scientific and a leisure one and offers two types of products for each of them. The goal of this project is to give the bookshop a way to represent the information about its products in XML and to verify that all information is in accordance with the rules described in the assignment. In this report, all the hypotheses that have been made as well as several examples will be seen.

2 XML Schema

The representations in the appendix show the schema in two different formats, with a tree and with relations. There are some non-interesting hypotheses, such as when an element must be present at least once or not. The two formats clearly show them. Therefore, in this section will only be seen the hypotheses that are worth talking about.

2.1 XSD Version

In order for the schema to be supported by the maximum number of systems, it has been decided to use XSD version 1.0.

2.2 ISBN

The International Standard Book Number is used to give to each book a unique identifier. For the sake of simplicity, it has been decided to only support ISBN 13. ISBN is composed of 13 digits distributed in five parts separated by a hyphen, one of which is a checksum digit. Due to the fact that XSD version 1.0 is used, the checksum digit cannot be checked, only the format is.

2.3 Currency

It has been decided for the element price to have an attribute currency that shows in which currency the price is. The currency is in its ISO 4217 code format and for the sake of simplicity only the following currencies are supported : EUR for Euro, USD for United States dollar, GBP for Pound sterling, JPY for Japanese yen.

2.4 Genre

The genre is an attribute that can take the following values : thriller, horror, sci-fi, romance and literature.

2.5 Price

It has been decided that the price should have as a value an integer or a decimal number that is greater or equal to 0. The decimal number is represented with a dot between the whole number part and the

fractional part.

2.6 Volume

It has been decided, for the sake of simplicity that the volume is identified with Arabic numerals.

3 Miscellaneous

It was wanted not to constrain the order in which the elements are placed as long as they are present. But in a special case a problem arose, when an element can be replaced by another, but both cannot be present at the same time. The naivest way to achieve this was to put the two elements in a **choice** in an **all** but XSD does not allow that. So, in order to do that it was needed to create a dummy abstract element and to include the two elements in the substitution group of the dummy element.

4 XML Examples

Several examples that are valid with respect to the XSD have been submitted. Each of these examples represents a different case from the other examples. The description of each example is given below.

Examples 1, 2 and 3 show that a bookshop can have only the science department. On the other hand, examples 4, 5 and 6 represent a bookshop where only the leisure department is present. However, example 7 contains both departments.

- Example 1 : This example contains one type of product from the science department which is the book. The format of a book can take many forms. Thus, the example contains 5 books each representing a possible form:
 - Book 1 : This format represents a book that contains only the mandatory information.
 - Book 2 : This format represents a book that contains the mandatory information and some optional information.
 - Book 3 : This format represents a book that contains a list of editors.
 - Book 4 : This format represents a book that contains a list of authors.
 - Book 5 : This format represents a book that contains the mandatory information and all optional information.
- Example 2 : This example contains the journal product of the scientific department. Two journal formats are given:
 - Journal 1 : This format represents a journal that contains mandatory information and optional information about the impact factor and article number.
 - Journal 2 : This format represents a journal that contains only the mandatory information.

- Example 3 : This example combines examples 1 and 2 to have a bookshop containing a scientific department that contains all its product types.
- Example 4 : This example contains one type of product from the leisure department which is the book. The format of a book can take many forms. Thus, the example contains some books possible form:
 - Book 1 : This format represents a book that contains the mandatory information and all optional information.
 - Book 2 and 3 : This format represents a book that contains the mandatory information with some possible values for its genre and some optional information.
 - Book 4 : This format represents a book that contains only the mandatory information.
- Example 5 : This example contains the periodical product of the leisure department. Only one format is approved, and it contains only the mandatory information.
- Example 6 : This example merges examples 4 and 5 to have a bookshop with a leisure department that contains all its product types.
- Example 7 : Here is a complete example containing both departments with the possible formats explained above for all different products.

A Relations

- $\text{bookshopType} \rightarrow (\text{scientificDepartment}, \text{scientificType})?, (\text{leisureDepartment}, \text{leisureType})?$
- $\text{scientificType} \rightarrow (\text{book}, \text{scientificBookType})^*, (\text{journal}, \text{scientificJournalType})^*$
- $\text{leisureType} \rightarrow (\text{book}, \text{leisureBookType})^*, (\text{periodical}, \text{leisurePeriodicalType})^*$
- $\text{scientificBookType} \rightarrow (\text{title}, \text{string}), (\text{authorsList}, \text{authorsListType}) \mid (\text{editorsList}, \text{editorsListType}), (\text{publisher}, \text{string}), (\text{year}, \text{nonNegativeInteger}), (\text{abstract}, \text{string})?, (\text{edition}, \text{string})?, (\text{ISBN}, \text{ISBN})?$
- $\text{scientificJournalType} \rightarrow (\text{title}, \text{string}), (\text{volume}, \text{nonNegativeInteger}), (\text{number}, \text{nonNegativeInteger}), (\text{authorsList}, \text{authorsListType}) \mid (\text{editorsList}, \text{editorsListType}), (\text{year}, \text{nonNegativeInteger}), (\text{publisher}, \text{string})?, (\text{impactFactor}, \text{impactFactorType})?, (\text{tableOfContents}, \text{tableOfContentsType})$
- $\text{leisureBookType} \rightarrow (\text{title}, \text{string}), (\text{authorsList}, \text{authorsListType}), (\text{publisher}, \text{string}), (\text{year}, \text{nonNegativeInteger}), (\text{genre}, \text{genreType}), (\text{edition}, \text{string})?, (\text{pagesNumber}, \text{nonNegativeInteger})?$
- $\text{leisurePeriodicalType} \rightarrow (\text{title}, \text{string}), (\text{price}, \text{priceType}), (\text{publisher}, \text{string})$
- $\text{authorsListType} \rightarrow (\text{author}, \text{string})^+$
- $\text{editorsListType} \rightarrow (\text{editor}, \text{string})^+$

- $\text{impactFactorType} \rightarrow (\text{year}, \text{nonNegativeInteger})$
- $\text{tableOfContentsType} \rightarrow (\text{article}, \text{articleType})^+$
- $\text{priceType} \rightarrow (\text{currency}, \text{currencyType})$
- $\text{articleType} \rightarrow (\text{title}, \text{string}), (\text{authorsList}, \text{authorsListType}), (\text{pair}, \text{pairType}) \mid (\text{number}, \text{nonNegativeInteger})$
- $\text{pairType} \rightarrow (\text{start}, \text{nonNegativeInteger}), (\text{end}, \text{nonNegativeInteger})$

B Tree

The blue rectangles mean a choice between two elements, in other words one of the two should be present but not both.

