**Софийски университет "Св. Климент Охридски", Факултет по математика и информатика**

2014 г.

**B2B XML transformation by SAX**

**Курсов проект по Маркъп езици за семантичен Уеб**

**Георги Антонов ф.н.: 61499, Явор Михайлов ф.н.: 61528**

# Задание

Компания А и компания Б използват различни XML документи за съхранение на техните продукти, които са подредени в подобни категории и имат подобни параметри. Да се направят два dtd файла, описващи семантиката на XML файловете описващи техните продукти. Единят файл трябва да е фокусиран върху употребата на атрибути, а другият върху елементи.

Да се направи приложение използващо SAX, което трансформира файловете на компания А към тези на Б и обратното, като запазва семантиката на информацията за продуктите.

# Информация за продуктите

Двете компании съхраняват следната информаця за продуктите:

* идентификационен код
* дата на добавяне
* категория
* производител
* модел
* описание
* свойства
* размери
* тегло
* цена

# Структура на XML документите

## Структура на XML документаите за компания A

|  |  |  |
| --- | --- | --- |
| Елемент | Описаие | Наследници |
| products | Корен на документа | product |
| product | Съдържа описание на отделните продукти | added-on, category, producer, model , description, features, dimensions, price |
| added-on | Дата на добавяне |  |
| category | Категория |  |
| producer | Производител |  |
| model | Модел |  |
| description | Описание на продукта в свободн текст. Не е задължителен. |  |
| features | Съдържа информаиция за свойствата на продукта . Не е задължителен. | feature |
| feature | Съдържа стойността на дадено свойство | feature-name, feature-value |
| feature-name | Име на свойството |  |
| feature-value | Стойност на свойството |  |
| dimensions | Информация за размерите. Не е задължителен. | length, width, height, weight |
| length, width, height, weigh | Описват размерите и теглото на профукта | unit, value |
| unit | Мерна единица |  |
| value | стойност |  |
| price | Цената на продукта | currency, cost |
| currency | Валутата, в която е зададена цената |  |
| cost | Цената зададена с число |  |

## Структура на XML документаите за компания Б

|  |  |  |  |
| --- | --- | --- | --- |
| Елемент | Описаие | Наследници | Атрибути |
| products | Корен на документа | product |  |
| product | Съдържа описание на отделните продукти | description, features, dimensions, price | Id, added-on, category, producer, model |
| description | Описание на продукта в свободн текст. Не е задължителен. |  |  |
| features | Съдържа информаиция за свойствата на продукта . Не е задължителен. | feature |  |
| feature | Съдържа стойността на дадено свойство |  | name - името на свойството |
| dimensions | Информация за размерите. Не е задължителен. | length, width, height, weight |  |
| length, width, height, weigh | Описват размерите и теглото на профукта |  | Unit |
| Price | Цената на продукта |  | currency |

# Реализация на приложението

Приложението приема от конзолата пътя до изходния файл, пътя до крайния файл и типа на изходния файл. Валидира изходния файл чрез посоченото в него dtd и преобразува в алтернативния тип файл.

За целта се използват два SAX Handler класа, единият от които трансформира тип А в тип Б, а другият - тип Б в тип А.