

Laboratorio 2. XML e DTD con DOM

In questo laboratorio verranno messi in pratica alcuni concetti teorici relativi ai documenti XML e alla loro elaborazione sfruttando l'interfaccia di programmazione DOM. In particolare verrà illustrato come costruire un DTD che rispetti i requisiti e come scrivere un documento valido e ben formato che la rispetti. Inoltre sarà usata l'interfaccia di programmazione per verificare la correttezza del documento e per estrarne delle informazioni.

Il DTD e i documenti XML definiti in questo laboratorio hanno lo scopo di rappresentare una base di dati relativa ad un insieme non troppo ampio di CD-ROM mantenendo le informazioni descritte di seguito.

Per ogni CD-ROM vengono specificate alcune informazioni: il titolo (obbligatorio), almeno un autore, il prezzo (facoltativo), delle note (facoltative) ed il tipo (facoltativo) che può essere ISO (dati) o AUDIO (musicale). Per ogni autore deve essere specificato obbligatoriamente il nome. Il tipo ISO è caratterizzato da tre proprietà: il tipo di ISO (ISO9660 o Joilet), il tipo di scrittura (multisessione o unica sessione) e la dimensione in MB. I valori predefiniti di queste proprietà sono rispettivamente ISO9660, unica sessione e 0. Il tipo AUDIO è caratterizzato da due attributi facoltativi: numero tracce audio e durata totale in minuti. Ogni CD-ROM è caratterizzato da un identificatore unico e l'ordine con cui devono comparire le informazioni per ognuno di essi è: titolo, autori, tipo CD, prezzo e note.

Suggerimenti

1. Provate prima a definire un file XML come descritto a parole e poi a scriverne il DTD.
2. Ricordate gli operatori di specifica degli elementi nel DTD:
 - , sequenza
 - | opzione
 - * zero o più
 - + uno o più
 - ? zero o uno
 - #REQUIRED attributo obbligatorio
 - #IMPLIED attributo opzionale
 - ecc... (vedi dispense lezione)
3. Ricordate la sequenza di creazione di un DOM:
 - Recuperare l'istanza del `DocumentBuilderFactory` e configurarlo
 - Usare la factory per generare un DOM del documento XML
 - Effettuare il parsing
 - Usare il DOM per stampare alcune informazioni: document type, numero di cd, titolo e prezzo di ogni cd, prezzo totale della collezione

Le classi con cui avrete a che fare sono `DocumentBuilderFactory`, `DocumentBuilder`, `Document`, `Element`, `Node` e `NodeList` del package `org.w3c.dom` (vedi documentazione).

Inoltre, è consigliabile procedere per gradi nella seguente maniera:

1. Scrivere il documento XML e il DTD
2. Scrivere il codice per validare il documento XML **senza** DTD (cioè, controllare se il documento è well-formed)
3. Aggiungere il DTD al documento XML e validarlo
4. Aggiungere il codice per l'analisi del contenuto del file XML