

UE Techniques avancées de gestion des données

AA Administration de BD et XML

3ième Bachelier en informatique de gestion (230x)
Haute Ecole de la Province de Liège (HEPL)

Ludovic Kutu <ludovic.kutu@hepl.be>
Laurence Herbiet <laurence.herbiet@hepl.be>

2015 – 2016

Introduction

Ce travail constitue la première partie de ce qui est demandé dans l'activité d'apprentissage (AA) "Administration de BD et XML" de l'unité d'enseignement (UE) "Techniques avancées de gestion des données" de la section 3ième bachelier en informatique de gestion. Il s'agit des technologies liées à XML qu'on ne sait pas évaluer de manière intégrée avec l'AA "Organisation et exploitation des données 4".

Document XML simple - MCD-SMALL-DOC

Cette fonctionnalité consiste à écrire un document XML permettant de représenter un schéma conceptuel particulier exprimé à l'aide du modèle entité-association (EA). Un schéma conceptuel DB-Main nommé `mcd_small.lun` vous est fourni (cfr. figure 1). On vous demande de le représenter.

Créez un autre document XML représentant un schéma entité-association.

Dans le cadre de cet exemple, on manipule un nombre limité d'objets (pas dans le sens OO du terme) du modèle EA qui sont des entités (uniquement fortes), des associations, des attributs (optionnels, obligatoires), des rôles et leurs cardinalités et des identifiants (uniquement internes).

On ne représente pas les types de données.

DTD simple - MCD-SMALL-DTD

Ecrivez un DTD permettant de valider les document XML créés précédemment. Effectuez les validations manuellement.

FIG. 1: Schéma entité-association `mcd_small.lun`

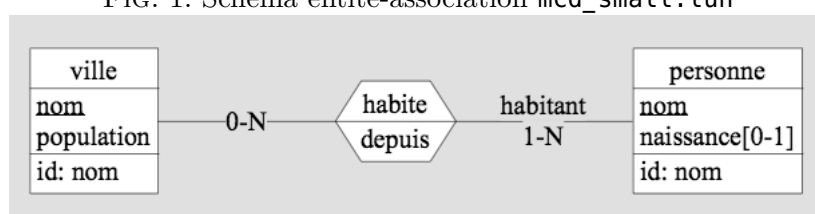
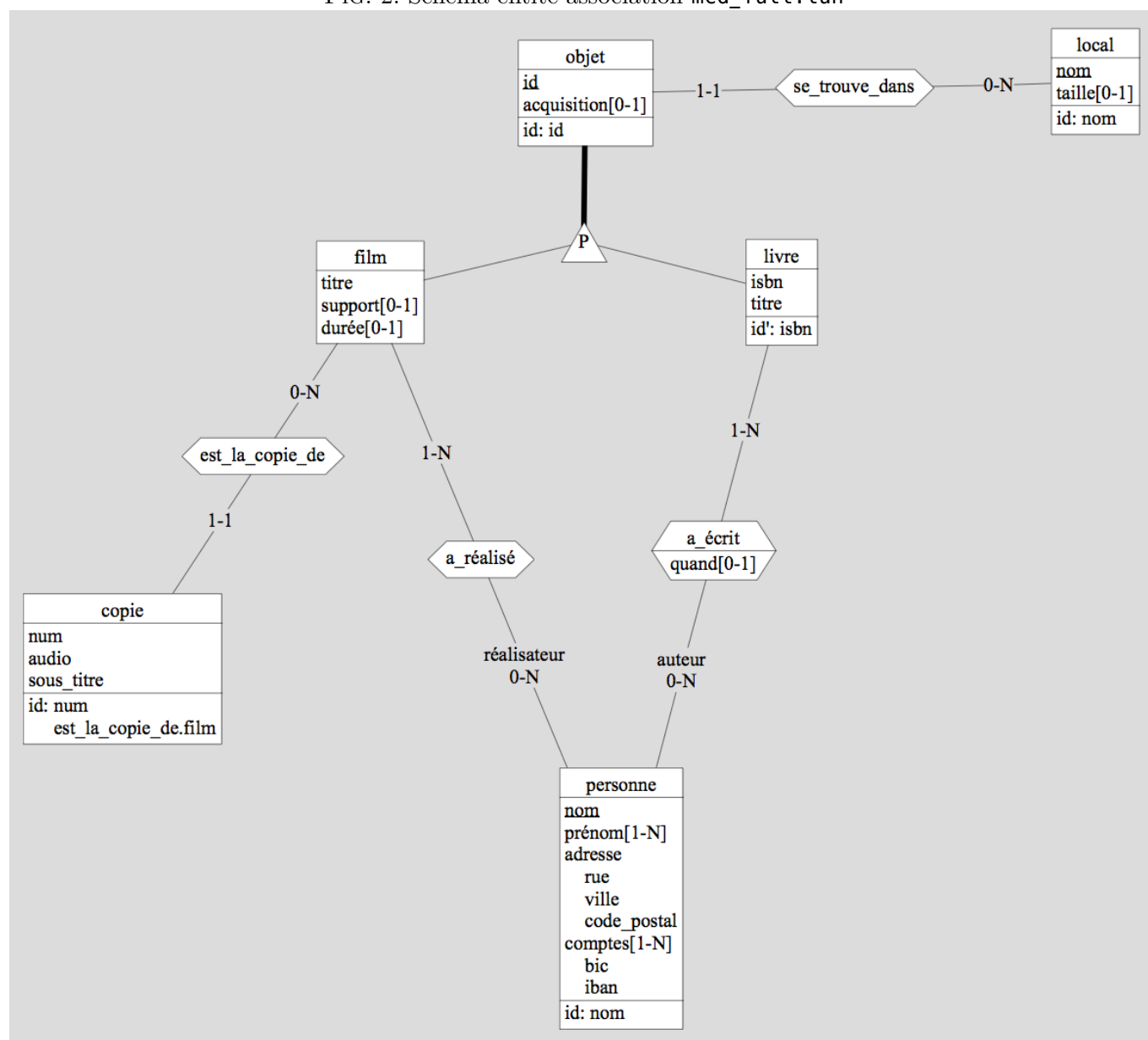


FIG. 2: Schéma entité-association `mcd_full.lun`

Espace de noms - MCD-SMALL-NS

Vous allez choisir un espace de noms à utiliser pour vos documents XML. De cette manière vous définissez un vocabulaire XML utilisable dans différents contextes. Ce vocabulaire est nommé MCD-SMALL. Adaptez les documents XML créés précédemment et le DTD pour qu'ils utilisent l'espace de noms choisi et qu'on puisse les valider manuellement.

Vocabulaire XML complet - MCD-FULL

Cette fonctionnalité consiste à augmenter le vocabulaire XML précédent pour permettre la représentation de tous les objets du modèle EA cités ci-dessous. Le nouveau vocabulaire est appelé MCD-FULL. Un schéma conceptuel DB-Main plus complet nommé `mcd_full.lun` vous est fourni (cfr. figure 2). On vous demande de le représenter et de le valider.

Il vous est donc demandé d'augmenter le DTD et faire un document XML de test basé sur `mcd_full.lun` pour tester la validation. L'espace de noms reste le même.

On manipule tous les objets du modèle EA qui sont des entités fortes et faibles, des associations, des attributs (optionnels, obligatoires, composés et/ou multivalués), des rôles et leurs cardinalités, des identifiants (primaires, secondaires, internes, externes), des hiérarchies, etc. Veillez à permettre la modélisation dans votre vocabulaire XML de tous ces concepts.

On ne représente pas les types de données.

Affichage des documents XML FULL - MCD-SAX

Ecrivez un programme Java sans GUI qui permet de faire une sérialisation de type `plain/text` claire et complète d'un document XML respectant le vocabulaire MCD-FULL vers un fichier texte qui portera le même nom que votre fichier XML mais en changeant l'extension en `txt`.

Le format de sortie contiendra toutes les informations présentes dans le document XML. Un exemple de sérialisation est donné au listing 1.

Listing 1: Exemple de sortie du programme SAX

```
SCHEMA(  
  entities:  
    ENTITY(  
      name:livre  
      attributes:  
        ATTRIBUTE(isbn)  
        ATTRIBUTE(titre)  
      identifiants:  
        IDENTIFIER(isbn)  
    )  
    ENTITY(  
      name:copie  
      attributes:  
        ATTRIBUTE(num)  
        ATTRIBUTE(audio)  
        ATTRIBUTE(subtitle)  
      identifiants:  
        IDENTIFIER(  
          num  
          est_la_copie_de.film  
        )  
    )  
    ...  
  relationships:  
    RELATIONSHIP(  
      name:a_écrit  
      attributes:  
        ATTRIBUTE(quand)  
      roles:  
        ROLE(1-N,a_écrit <-> livre)  
        ROLE(auteur,0-N,a_écrit <-> personne)  
    )  
    ...  
  generalizations:  
    GENERALIZATION(  
      total:true  
      disjoint:true
```

```
    supertype:objet
    subtypes:
      livre
      film
  )
)
```

Transformation des entités - MCD-DOM

A l'aide de DOM 2, vous allez transformer les documents XML MCD-FULL en effectuant les opérations suivantes dans l'ordre :

- Un attribut univalué composé est remplacé par des attributs avec un nouvel attribut par sous-attribut de l'attribut original dont le nom est la concaténation du nom de l'attribut original avec le nom du sous-attribut séparé par un underscore. Dans notre cas, **adresse** devient **adresse_rue**, **adresse_ville** et **adresse_code_postal**.
- Un attribut multivalué simple ou composé est remplacé par une entité qui aura comme nom la concaténation du nom de l'entité originale avec le nom de l'attribut séparé par un underscore. Pensez à créer un identifiant correct et une association avec les bonnes cardinalités. Veillez aussi à créer les attributs nécessaires dans la nouvelle entité.

On suppose que les sous-attributs d'un attribut composé (top-level) ne peuvent être que simples et univalués. S'ils ne le sont pas, on ne transforme pas l'attribut parent.

Sérialisez l'arbre manuellement (pas de transformation XSLT identité, pas de DOM 3, pas de sérialisation obsolète de Xerces) dans un fichier XML dont le nom sera **monfichier_mod.xml** si le fichier source se nomme **monfichier.xml**. Utilisez ensuite la fonctionnalité MCD-SAX (section , page 3) pour effectuer une sérialisation de **monfichier_mod.xml**. On obtiendra donc un fichier nommé **monfichier_mod.txt**.