

TECHNIQUES DE MODÉLISATION

Dictionnaire de données

TM040

Luc LAVOIE, Christina KHNAISSER
Département d'informatique
Faculté des sciences



UNIVERSITÉ DE
SHERBROOKE

Luc.Lavoie@USherbrooke.ca
<http://info.usherbrooke.ca/llavoie>

TABLE DES MATIÈRES

- Présentation
- Types
- Importance
- Contenu
- Notation
- Implémentation
- Vocabulaire usuel
- Références
- À suivre

PRÉSENTATION

- Outils de modélisation, indispensable.
- C'est une liste qui décrit de toutes les données en circulation dans le logiciel avec précision de telle sorte qu'elle soit compréhensible pour l'utilisateur et l'équipe de développement.

TYPES

- Textuel (glossaire)
- Formel :
 - BNF (*Backus Naur Form*),
 - GC (Grammaire de Constituants)
 - GA (Grammaire Attribuée).

IMPORTANCE

- Permet aux membres de l'équipe d'avoir accès à la même information.
- Vérification de la compréhension des détails de l'application

CONTENU

Description du diagramme de flux de donnée :

- Flux de données
- Dépôts de données
- Acteurs

Description du modèle conceptuel de données :

- Les entités et leurs attributs
- Les relations entre entités

NOTATION

Symbole	Description	Exemple
=	Définit un élément (est composé de)	$a = b$
+	Et	$c + d$
()	Élément optionnel	$a (+b)$
{ }	Itération	$s = \{ c \}$
[]	Sélection d'un ou plusieurs choix	$[a \mid b \mid c]$
@	Identificateur de stockage	@a
	Ou	$a \mid c$
**	Commentaires	** Note

IMPLÉMENTATION

- Certains standards exigent de distinguer explicitement :
 - la description en en relation avec le problème,
 - le domaine des valeurs,
 - la représentation externe.
- Ceci facilite la revue au détriment, parfois, de la longueur des définitions.

○ Édition

- L'édition correspond à une variation d'un ouvrage comportant des changements quant à la forme ou au contenu. Elle se distingue de la réimpression qui est une reproduction à l'identique. Une édition comporte un numéro et une mention (par exemple « revue », « revue et corrigée », « augmentée », « revue et augmentée », etc.). Dans le cadre de notre système, il a été décidé de ne pas représenter la mention de l'édition. Une édition est donc représentée par entier compris entre 1 et 99.

○ Groupe

- Instance particulière d'une activité pédagogique au sein d'un trimestre ; entier compris entre 0 et 99.

○ Id_Livre

- Identifiant interne unique d'un livre ; un entier compris entre 1 000 000 et 9 999 999.

○ ISBN

- Il existe deux formats d'ISBN, un à dix symboles (ancien) et un autre à treize symboles (nouveau) ; par convention, les trois premières positions des anciens ISBN prennent pour valeur l'espace. Les symboles sont des chiffres décimaux, sauf pour la dernière position ; celle-ci peut être un chiffre ou la lettre `X` (majuscule). Des tirets intercalaires sont parfois insérés pour faciliter la lecture, ils n'ont pas de signification et peuvent être ignorés.

ISBN ::= ISBN10 | ISBN13

ISBN10 ::= { ` ` }³ { chiffreI }⁹ chiffreX

chiffre ::= `0` | `1` | `2` | ... | `9`

chiffreI ::= chiffre ['-']

chiffreX ::= chiffre | `X`

ISBN13 ::= { chiffreI }¹² chiffreX

○ Prix

- Prix d'un bien échangeable (pour le moment un livre) ; le prix est exprimé en CAD à deux décimales près ; il est inférieur à 10 000 CAD.

Prix ::= { chiffre }⁴ ` , ` { chiffre }²

○ Texte(n)

- Texte encodé dans l'un des trois alphabets occidentaux suivants : latin, grec, cyrillique ; peut comprendre, ponctuation, chiffres, etc. ; longueur maximale n caractères ; si le texte est effectivement plus long, il est tronqué à n-1 caractères et le dernier caractère prend la valeur `...` (trois points de suspension).

○ Trimestre

- L'année universitaire est divisée en trois périodes appelées « trimestres » (sic) :
Trimestre ::= Automne | Hiver | Été

Domaine	Définition
Édition	L'édition correspond à une variation d'un ouvrage comportant des changements quant à la forme ou au contenu. Elle se distingue de la réimpression qui est une reproduction à l'identique. Une édition comporte un numéro et une mention (par exemple « revue », « revue et corrigée », « augmentée », « revue et augmentée », etc.). Dans le cadre de notre système, il a été décidé de ne pas représenter la mention de l'édition. Une édition est donc représentée par entier compris entre 1 et 99.
Groupe	Instance particulière d'une activité pédagogique au sein d'un trimestre ; entier compris entre 0 et 99.
Id_Livre	Identifiant interne unique d'un livre ; un entier compris entre 1 000 000 et 9 999 999.
ISBN	<p>Il existe deux formats d'ISBN, un à dix symboles (ancien) et un autre à treize symboles (nouveau) ; par convention, les trois premières positions des anciens ISBN prennent pour valeur l'espace. Les symboles sont des chiffres décimaux, sauf pour la dernière position ; celle-ci peut être un chiffre ou la lettre 'X' (majuscule). Des tirets intercalaires sont parfois insérés pour faciliter la lecture, ils n'ont pas de signification et peuvent être ignorés.</p> <pre> ISBN ::= ISBN10 ISBN13 ISBN10 ::= { ` ` }³ { chiffreI }⁹ chiffreX chiffre ::= `0` `1` `2` ... `9` chiffreI ::= chiffre ['-'] chiffreX ::= chiffre `X` ISBN13 ::= { chiffreI }¹² chiffreX </pre>
Prix	<p>Prix d'un bien échangeable (pour le moment un livre) ; le prix est exprimé en CAD à deux décimales près ; il est inférieur à 10 000 CAD.</p> <pre>Prix ::= { chiffre }⁴ `.` { chiffre }²</pre>
Ref_Livre	Identifiant externe unique d'un livre ; son format est prescrit par le regroupement d'éditeurs fournissant les descriptions de livres et leur mise à jour, il varie donc fréquemment. Il a été convenu d'en limiter la longueur à 32 caractères, les caractères étant eux-mêmes limités aux lettres latines non accentuées, chiffres, espace et tiret. La distinction entre majuscule et minuscule doit être conservée à l'affichage, bien qu'elle ne soit pas significative par ailleurs.
Texte(n)	Texte encodé dans l'un des trois alphabets occidentaux suivants : latin, grec, cyrillique ; peut comprendre, ponctuation, chiffres, etc. ; longueur maximale n caractères ; si le texte est effectivement plus long, il est tronqué à n-1 caractères et le dernier caractère prend la valeur `...` (trois points de suspension).
Trimestre	<p>L'année universitaire est divisée en trois périodes appelées « trimestres » (sic) :</p> <pre>Trimestre ::= Automne Hiver Été</pre>

- Noter la différence entre les définitions utilisant des grammaires (de type BNF ici) et celles reposant uniquement sur une description textuelle.

EXEMPLE TABULAIRE

FOURNI PAR MICHEL RONDEAU (2015)

Attribut	Description	Type (Forme)							Contrainte	Domaine de valeur	Informations complémentaires	Entité
		Simple	Composite	Mono-valué	Multi -valué	Stocké	Dérivé	Null				
No_Appareil	Le N° de l'appareil d'éclairage physique	X		X		X			Identifiant unique			Appareil

- Il est possible de réduire le nombre de colonnes dans le groupe "Type". En effet, simple et composite sont mutuellement exclusifs (on pourrait d'ailleurs reporter cette information dans la définition du domaine de valeurs).
- Mono et multi-valué sont aussi mutuellement exclusifs (mais ne font pas nécessairement partie du domaine de valeur, c'est pourquoi plusieurs préfèrent, comme ici, l'indiquer dans la description de l'attribut).
- Stocké, n'a pas d'intérêt au niveau de l'analyse (il en a au niveau de la conception).
- Dérivé ne devrait pas figurer seul (la formule de dérivation, de calcul, DOIT être fournie).
- Il y a un débat entre les spécialistes à savoir si NULL doit faire partie du domaine de valeurs ou de la définition de l'attribut : choisissez votre camp et soyez systématiques!

- En fait, il y a un chevauchement intrinsèque entre le MCD et le DD.
- Certains auteurs (et praticiens) préfèrent maximiser l'information dans la description du MCD et de ses entités (le DD se résume alors le plus souvent en la définition des domaines, comme dans l'exemple que je vous ai soumis), d'autres préfèrent maximiser le DD (le diagramme du MCD est alors très épuré, ne comporte pas tous les attributs et présente, à l'aide de son texte d'accompagnement, une vue de haut niveau; ce qui n'empêche pas que les domaines doivent être décrits séparément... comme dans l'exemple que j'ai soumis).
- En conclusion, la définition des domaines au sein du DD est incontournable. La définition des attributs (voir des entités) peut être placée dans le MCD ou dans le DD. Il n'y a pas de "meilleure" méthode selon moi. Encore une fois, le critère prépondérant doit être la qualité de la validation que pourra en faire l'expert du domaine. Ensuite, on pourra considérer des critères comme les outils à disposition, le nombre de répétitions (lorsqu'il y a plusieurs entités semblables), etc.

VOCABULAIRE USUEL



RÉFÉRENCES

- Ed Yourdon, 2006



À SUIVRE