

TP2-Introduction aux bases de données

Durée : 2h

1 Introduction

L'administration, les banques, les assurances, les secteurs de la finance utilisent des **bases de données**, systèmes d'informations qui stockent dans des fichiers les données nombreuses qui leur sont nécessaires.

Une base de données relationnelle permet d'organiser, de stocker, de mettre à jour et d'interroger des données structurées volumineuses utilisées simultanément par différents programmes ou différents utilisateurs. Un logiciel, le **système de gestion de bases de données** (SGBD), est utilisé pour la gestion (lecture, écriture, cohérence, actualisation ...) des fichiers dans lesquels sont stockées les données.

L'accès aux données d'une base de données relationnelle s'effectue en utilisant un langage informatique qui permet de sélectionner des données spécifiées par des formules de logique, appelées **requêtes** d'interrogation et de mise à jour.

L'objectif est de ce TP est de présenter le vocabulaire des bases de données. Un TP ultérieur aura pour but de présenter une description applicative des bases de données en langage de requêtes **SQL** (*Structured Query Language*).

2 Bases de données et SGBD

Base de données

Une base de données est un ensemble d'informations structurées mémorisées sur un support de stockage (disque dur par exemple).

Exemple 1

1. Les fichiers clients d'une entreprise forment une base de données recensant les informations relatives aux clients de l'entreprise (coordonnées, l'historique de leurs achats,...).
2. Les fichiers de gestion de stocks forment une base de données recensant les informations relatives aux marchandises d'une entreprise (nom des produits, état du stock, achats, ventes,...).

Dans une base de données, les utilisateurs n'ont pas un accès direct aux fichiers constituant la base de données. En effet, si c'était le cas, il y aurait des risques de suppression ou d'altération des fichiers par erreur humaine (intégrité), il pourrait survenir des problèmes liés à l'accès simultané de plusieurs utilisateurs à un même fichier, ...

Système de Gestion de Bases de Données

Un Système de Gestion de Bases de Données est un système informatique qui assure la gestion de l'ensemble des informations stockées dans la base. Il a notamment pour rôle :

- d'éviter les risques évoqués ci-dessus,
- de gérer l'interaction avec les utilisateurs.

3 Vocabulaire des bases de données

3.1 Modèle relationnel

Comme les utilisateurs n'ont pas un accès direct aux fichiers, le système de gestion de bases de données peut choisir un mode de représentation des données (table, graphe, liste ...) indépendamment de la façon dont les données sont codées sur le disque.

Un **modèle** est la donnée d'un mode représentation des données ainsi que d'un langage pour interagir avec les données.

Modèle relationnel

Le **modèle relationnel** est caractérisé par

- un mode de représentation des données par des **tables** (tableaux) dont chaque ligne représente un objet (et tous les objets sont de même nature),
- le langage SQL pour interagir avec les données.

Une base de données est alors formée d'une ou plusieurs tables.

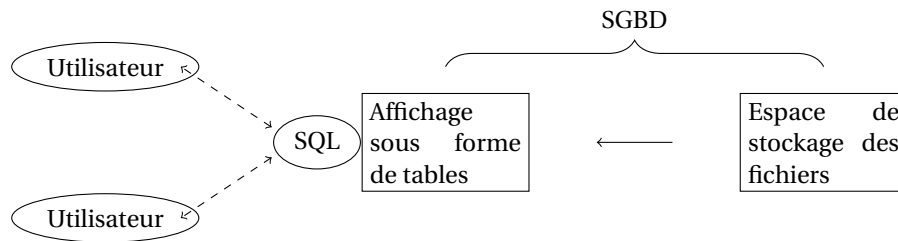


FIGURE 1 – Modèle relationnel

Exemple 2

Un extrait d'une base de données de gestion de stock d'une librairie :

Titre	Auteur	Année de publication	Genre	Quantité en stock
<i>Les Fleurs du Mal</i>	<i>Baudelaire</i>	1857	<i>Poésie</i>	3
<i>Le Spleen de Paris</i>	<i>Baudelaire</i>	1869	<i>Poésie</i>	5
<i>Le Rouge et le Noir</i>	<i>Stendhal</i>	1830	<i>Littérature française</i>	1
<i>Madame Bovary</i>	<i>Flaubert</i>	1857	<i>Littérature française</i>	1
<i>Faust I</i>	<i>Goethe</i>	1808	<i>Théâtre</i>	2
<i>Les raisins de la colère</i>	<i>Steinbeck</i>	1939	<i>Littérature américaine</i>	2
<i>Joyland</i>	<i>Schultz</i>	2005	<i>Littérature canadienne</i>	0
<i>Joyland</i>	<i>King</i>	2013	<i>Littérature américaine</i>	3

Remarque 1

Dans une table, l'ordre des lignes n'a pas d'importance.

Vocabulaire

Dans une même **table** on regroupe des éléments d'une même entité (clients, objets, ...) sur lesquelles on dispose d'un certain nombre d'informations.

- Chaque **ligne** de la table représente les informations associées à un élément. Une ligne est aussi appelée un **enregistrement**.
- Chaque information dont on dispose est appelée un **attribut** : chaque colonne de la table correspond à un attribut.
- Le type d'un attribut est appelé son **domaine**. On rencontrera deux types : « INT » (nombre entier) et « TEXT » (texte).
- L'ensemble des données regroupées dans une table est appelée une **relation** (table et relation sont synonymes).

Une table dont le nom est NOM_TABLE ayant pour attributs Attribut_1,...,Attribut_n de domaine respectif Domaine_1,..., Domaine_n est notée :

NOM_TABLE(Attribut_1:Domaine_1,...,Attribut_n:Domaine_n).

Cette notation est appelée le **schéma de relation**.

Exemple 3

On reprend l'exemple d'une base de données de gestion de stock d'une librairie :

Titre	Auteur	Année de publication	Genre	Quantité en stock
<i>Les Fleurs du Mal</i>	<i>Baudelaire</i>	1857	<i>Poésie</i>	3
<i>Le Spleen de Paris</i>	<i>Baudelaire</i>	1869	<i>Poésie</i>	5
<i>Le Rouge et le Noir</i>	<i>Stendhal</i>	1830	<i>Littérature française</i>	1
<i>Madame Bovary</i>	<i>Flaubert</i>	1857	<i>Littérature française</i>	1
<i>Faust I</i>	<i>Goethe</i>	1808	<i>Théâtre</i>	2
<i>Les raisins de la colère</i>	<i>Steinbeck</i>	1939	<i>Littérature américaine</i>	2
<i>Joyland</i>	<i>Schultz</i>	2005	<i>Littérature canadienne</i>	0
<i>Joyland</i>	<i>King</i>	2013	<i>Littérature américaine</i>	3

1. Ici, chaque ligne correspond à un élément : un livre.
2. Chaque colonne correspond à un attribut : **Titre**, **Auteur**, **Année de publication**, **Genre**, **Quantité en stock**.
3. Les attributs **Titre**, **Auteur** et **Genre** sont de type « TEXT ».
4. Les attributs **Année de publication**, **Quantité en stock** sont de type « INT ».
5. Le schéma de la relation s'écrit :

Livres(Titre: TEXT,Auteur: TEXT,Année de publication: INT,Genre: TEXT,Quantité en stock: INT).

Souvent pour le nom de la table on prend le nom des entités qu'elle représente : ici des livres.

On considère la relation Salles suivante :

Numéro	Bâtiment	Capacité	Vidéoprojecteur	Accès
104	C	40	Oui	Clé
107	A	35	Non	Badge
302	C	60	Oui	Clé
201	B	35	Oui	Clé

- Donner les attributs de la relation *Salles* et leur domaine respectif.

- Combien cette relation compte-t-elle d'enregistrements?

- Écrire le schéma de la relation.

3.2 Clés primaires

Si l'on reprend l'exemple du stock de livres, il est tout à fait possible que deux livres aient exactement le même titre ou le même auteur ... Dans une table, il est très important d'identifier de manière **unique** chaque ligne. C'est pour résoudre ce problème d'identification qu'interviennent les clés.

Clé primaire

Une **clé primaire** d'une table est un groupe minimal d'attributs qui permet d'identifier de manière unique les enregistrements de la table : deux enregistrements distincts ont deux clés primaires distinctes.

Exemple 4

On reprend la table de l'exemple précédent :

<i>Titre</i>	<i>Auteur</i>	<i>Année de publication</i>	<i>Genre</i>	<i>Quantité en stock</i>
<i>Les Fleurs du Mal</i>	<i>Baudelaire</i>	1857	<i>Poésie</i>	3
<i>Le Spleen de Paris</i>	<i>Baudelaire</i>	1869	<i>Poésie</i>	5
<i>Le Rouge et le Noir</i>	<i>Stendhal</i>	1830	<i>Littérature française</i>	1
<i>Madame Bovary</i>	<i>Flaubert</i>	1857	<i>Littérature française</i>	1
<i>Faust I</i>	<i>Goethe</i>	1808	<i>Théâtre</i>	2
<i>Les raisins de la colère</i>	<i>Steinbeck</i>	1939	<i>Littérature américaine</i>	2
<i>Joyland</i>	<i>Schultz</i>	2005	<i>Littérature canadienne</i>	0
<i>Joyland</i>	<i>King</i>	2013	<i>Littérature américaine</i>	3

1. L'attribut **Auteur** ne peut pas être une clé primaire car deux enregistrements distincts ont la même valeur d'attribut : *Les Fleurs du Mal* et *Le Spleen de Paris*. De même l'attribut **Titre** n'est pas une clé primaire.
2. En revanche la paire {**Titre**, **Auteur**} est une clé primaire.

On considère la relation *Salles* suivante :

Numéro	Bâtiment	Capacité	Vidéoprojecteur	Accès
104	C	40	Oui	Clé
107	A	35	Non	Badge
302	C	60	Oui	Clé
201	B	35	Oui	Clé

- Pour chacun des attributs, dire s'il peut jouer le rôle de clé primaire en justifiant.

Remarque 2

Le choix d'une clé primaire pourrait très bien être arbitraire (parmi les clés candidates). Cependant, on choisit un général une clé primaire de façon à ce qu'elle soit simple, c'est-à-dire qu'elle ne contienne le moins d'attributs possibles et que ces attributs soient "basiques" (par exemple des entiers ou des chaînes de caractères courtes).

À cette fin, il est courant d'ajouter artificiellement un attribut (un numéro d'identification) à la relation et qui n'a pour seul but que d'identifier de manière unique chaque enregistrement.

Exemple 5

On reprend la table des exemples précédents à laquelle on ajoute un attribut :

<i>Id</i>	<i>Titre</i>	<i>Auteur</i>	<i>Année de publication</i>	<i>Genre</i>	<i>Quantité en stock</i>
1	<i>Les Fleurs du Mal</i>	<i>Baudelaire</i>	1857	<i>Poésie</i>	3
2	<i>Le Spleen de Paris</i>	<i>Baudelaire</i>	1869	<i>Poésie</i>	5
3	<i>Le Rouge et le Noir</i>	<i>Stendhal</i>	1830	<i>Littérature française</i>	1
4	<i>Madame Bovary</i>	<i>Flaubert</i>	1857	<i>Littérature française</i>	1
5	<i>Faust I</i>	<i>Goethe</i>	1808	<i>Théâtre</i>	2
6	<i>Les raisins de la colère</i>	<i>Steinbeck</i>	1939	<i>Littérature américaine</i>	2
7	<i>Joyland</i>	<i>Schultz</i>	2005	<i>Littérature canadienne</i>	0
8	<i>Joyland</i>	<i>King</i>	2013	<i>Littérature américaine</i>	3

L'attribut **Id** est une clé primaire.

3.3 Clés étrangères

En général, une base de données contient plusieurs tables et l'on peut souhaiter croiser les données présentes dans plusieurs d'entre elles. Dans l'exemple de la base de données de gestion des stocks de la librairie, nous pouvons nous intéresser aux commandes dont chaque livre fait l'objet. On est donc amené à lier la base de données de gestion de stocks avec la base de données des commandes :

Id_commande	Nom du client	Livre commandé
101	Dupond	4
102	Dupon	6
103	Dupont	2

Si on souhaite savoir quels livres sont commandés, on doit connaître leur **Id** puis aller chercher dans la table des commandes dans quelles lignes ils apparaissent.

Une **clé étrangère** d'une table est un attribut qui référence la clé primaire d'une autre table.

On considère une base de données de gestion de stock d'une librairie qui contient une table *Livres* et *Commandes* :

Livres					
<i>Id</i>	<i>Titre</i>	<i>Auteur</i>	<i>Année de publication</i>	<i>Genre</i>	<i>Quantité en stock</i>
1	<i>Les Fleurs du Mal</i>	<i>Baudelaire</i>	1857	<i>Poésie</i>	3
2	<i>Le Spleen de Paris</i>	<i>Baudelaire</i>	1869	<i>Poésie</i>	5
3	<i>Le Rouge et le Noir</i>	<i>Stendhal</i>	1830	<i>Littérature française</i>	1
4	<i>Madame Bovary</i>	<i>Flaubert</i>	1857	<i>Littérature française</i>	1
5	<i>Faust I</i>	<i>Goethe</i>	1808	<i>Théâtre</i>	2
6	<i>Les raisins de la colère</i>	<i>Steinbeck</i>	1939	<i>Littérature américaine</i>	2
7	<i>Joyland</i>	<i>Schultz</i>	2005	<i>Littérature canadienne</i>	0
8	<i>Joyland</i>	<i>King</i>	2013	<i>Littérature américaine</i>	3

L'attribut **Livre commandé** est donc une clé étrangère de la table *Commandes*.

Vocabulaire

On appelle **schéma relationnel** d'une base de données la donnée des schémas de relation de chaque table en précisant les clés primaires et étrangères.

En général, on souligne les clés primaires et on place un dièse (#) devant les clés étrangères.

Exemple 7

On reprend l'exemple précédent. Le schéma relationnel est :

Livres(Id: INT, Titre: TEXT, Auteur: TEXT, Année de publication: INT, Genre: TEXT, Quantité en stock: INT)

Commandes(Id_commande: INT, Nom du client: TEXT, #Livre commandé: INT)

4 Exercice

On considère les trois tables suivantes :

Chats			
Id_chat	Nom	Propriétaire	Pelage
C1	Aramis	H3	nb1
C2	Lambda	H3	cb1
C3	Gerda	H4	nb2
C4	Lili	H5	tm1
C5	Barbaque	H1	r1
C6	Reblochon	H2	nb2

Propriétaires			
Id_propriétaire	Nom	Prénom	Téléphone
H1	Ghilder	Mateo	0668543452
H2	Troccovee	Anne	0675463309
H3	Chasoeler	Philip	0667240908
H4	Bobont	Charline	0638899821
H5	Ricciardo	Louis	0619798669

Pelages		
Id_pelage	Poil	Couleurs
r1	court	roux
nb1	court	noir à tâches blanches
cb1	court	crème et blanc
nb2	court	blanc à tâches noires
tm1	mi-long	noir, brun et blanc

- Comment s'appelle le propriétaire de Gerda?

- Combien de chats possède Philip? Donner leur nom ainsi que leurs couleurs.

- Déterminer les clés primaires de chacune des tables.

- Déterminer les clés étrangères de la table Chats.

- Écrire le schéma relationnel associée à ces trois tables.