

Bases de données relationnelles

Introduction

Les bases de données sont une des grandes réussites de l'informatique du vingtième siècle au même titre qu'Internet ou le Web.

De quoi s'agit-il ?

Nous avons vu l'année dernière que lorsque les données deviennent volumineuses, nous sommes obligés d'utiliser des fichiers externes (par exemple des fichiers .csv)

On pouvait alors par programmation extraire des informations d'une copie de ce fichier, **chacun de son côté** ce qui implique que :

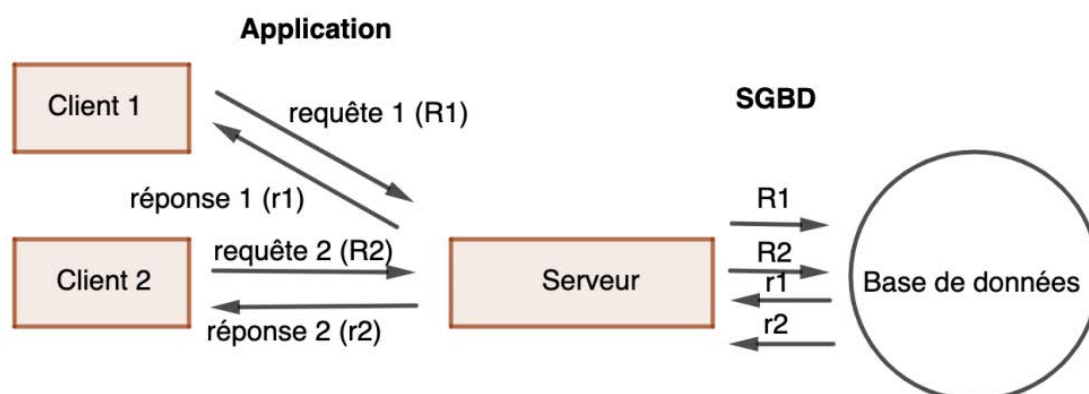
1. Les données sont **statiques** au sens où elles ne sont pas mises à jour régulièrement
2. Plusieurs personnes ne peuvent **pas consulter les mêmes données en même temps**

Or à l'heure actuelle il est courant que, des **données volumineuses et actualisées** soient consultées en ligne par plusieurs personnes, par exemple des réservations de billets de train ou de locations de vacances ou encore des cartes (OpenstreetMap).

On peut observer d'ailleurs cette "**concurrence**" dans la consultation des données grâce à des messages affichés sur la page Web comme "*il ne reste que 5 billets pour cette destination*" ou encore "*plusieurs personnes consultent cette destination*"

Schématiquement le principe est le suivant : plusieurs clients peuvent exprimer **des requêtes** en ligne à un serveur qui est en relation avec une base de données gérée par un logiciel appelé **système de gestion de bases de données (SGBD)**

<https://www.youtube.com/watch?v=8na73kwO-EI&t=1s>



Définition :

Une **base de données** est un ensemble structuré de données stockées sur un support persistant.
Un **système de gestion de base de données (SGBD)** est un logiciel qui gère les données de la base de données

Vocabulaire des bases de données

Au début des années 1970 à partir de travaux sur l'algèbre et la logique sont apparues les bases de données relationnelles. De manière intuitive une base de données relationnelles est un ensemble de tableaux à deux dimensions appelés **tables ou relations**.

Toutefois, commençons déjà par l'étude d'une seule table.

En voici un exemple :

Une table de livres

id	titre	auteur	ann_publi	note
1	1984	Orwell	1949	10
2	Dune	Herbert	1965	8
3	Fondation	Azimov	1951	9
4	Le meilleur des mondes	Huxley	1931	9
5	Fahrenheit 451	Bradbury	1953	7
6	Ubik	K.Dick	1969	6
7	Chroniques martiennes	Bradbury	1950	7
8	La nuit des temps	Barjavel	1968	7
9	Blade Runner	K.Dick	1968	9
10	Les Robots	Azimov	1950	10
11	La Planète des singes	Boulle	1963	9
12	Ravage	Barjavel	1943	11
13	Le Maître du Haut Château	K.Dick	1962	7
14	La fin de l'éternité	Azimov	1955	9

Table : LIVRES

Le vocabulaire est important dans le domaine des bases de données aussi on va préciser les termes employés :

Vocabulaire : https://www.youtube.com/watch?v=jDXU8PvsG3U&feature=emb_logo

Une table se décompose de deux parties :

- l'**en-tête** (première ligne)
- le **corps** (composé des autres lignes)

Un **enregistrement** ou un **t-uplet** est une ligne d'un tableau.

Un **descripteur** ou **attribut** correspond au titre d'une colonne d'un tableau.

Le **domaine** d'un attribut est l'ensemble des valeurs que l'attribut peut prendre.

Exemple :

Entête :

attributs :

Enregistrement 9 :

id	titre	auteur	ann_publi	note	en-tête
1	1984	Orwell	1949	10	
2	Dune	Herbert	1965	8	
3	Fondation	Azimov	1951	9	
4	Le meilleur des mondes	Huxley	1931	9	
5	Fahrenheit 451	Bradbury	1953	7	
6	Ubik	K.Dick	1969	6	
7	Chroniques martiennes	Bradbury	1950	7	
8	La nuit des temps	Barjavel	1968	7	
9	Blade Runner	K.Dick	1968	9	
10	Les Robots	Azimov	1950	10	un t-uplet
11	La Planète des singes	Boulle	1963	9	
12	Ravage	Barjavel	1943	11	
13	Le Maître du Haut Château	K.Dick	1962	7	
14	La fin de l'éternité	Azimov	1955	9	

attribut : titre corps

Clé primaire et clé étrangère

Dans une relation, deux t-uplets identiques sont **non autorisés** et il faut pouvoir **identifier de façon unique chaque t-uplet**.

Il faut déterminer, parmi les attributs, lequel permet d'identifier de manière unique un t-uplet. Cet attribut sera considéré comme la **clé primaire** de la relation.

Définition : <https://www.youtube.com/watch?v=ws7ihSHJDsg&t=214s>

Un attribut qui identifie de manière unique un n-uplet est appelé **clé primaire**.

Exemple :

On remarquera que *parfois une clé primaire est l'association de deux (ou plusieurs) attributs*.

D'autre part, il est très fréquent que l'on ajoute un attribut id pour jouer le rôle de cette clé primaire.

Ajoutons maintenant une deuxième relation à notre base de données :

id	titre	id_auteur	ann_publi	note
1	1984	1	1949	10
2	Dune	2	1965	8
3	Fondation	3	1951	9
4	Le meilleur des mondes	4	1931	9
5	Fahrenheit 451	5	1953	7
6	Ubik	6	1969	6
7	Chroniques martiennes	5	1950	7
8	La nuit des temps	7	1968	7
9	Blade Runner	6	1968	9
10	Les Robots	3	1950	10
11	La Planète des singes	8	1963	9
12	Ravage	7	1943	11
13	Le Maître du Haut Château	6	1962	7
14	La fin de l'éternité	3	1955	9

Relation : LIVRES

id	nom	prenom	ann_naissance	langue_écriture
1	Orwell	George	1903	anglais
2	Herbert	Frank	1920	anglais
3	Azimov	Isaac	1920	anglais
4	Huxley	Aldous	1894	anglais
5	Bradbury	Ray	1920	anglais
6	K. Dick	Philip	1928	anglais
7	Barjavel	René	1911	français
8	Boulle	Pierre	1912	français

Relation : AUTEURS

Lors de l'ajout de la nouvelle table, nous avons changé l'attribut *auteur* par l'attribut *id_auteur* permettant de faire le lien entre les deux relations.

id_auteur présente dans la relation **LIVRES** est la clé primaire d'une autre relation (**AUTEURS**) : on dira que c'est une **clé étrangère** de la relation **LIVRES**.

Définition : https://www.youtube.com/watch?v=5X_g8MAESjg

Un attribut d'une relation qui correspond à une clé primaire d'une autre relation est appelé **clé étrangère**.

Ce sont ces **clés étrangères** qui permettent de **relier** les différentes relations entre elles.

Clé primaire

Clé étrangère

Clé primaire

id	titre	id_auteur	ann_publi	note
1	1984	1	1949	10
2	Dune	2	1965	8
3	Fondation	3	1951	9
4	Le meilleur des mondes	4	1931	9
5	Fahrenheit 451	5	1953	7
6	Ubik	6	1969	6
7	Chroniques martiennes	5	1950	7
8	La nuit des temps	7	1968	7
9	Blade Runner	6	1968	9
10	Les Robots	3	1950	10
11	La Planète des singes	8	1963	9
12	Ravage	7	1943	11
13	Le Maître du Haut Château	6	1962	7
14	La fin de l'éternité	3	1955	9

Relation : LIVRES

id	nom	prenom	ann_naissance	langue_écriture
1	Orwell	George	1903	anglais
2	Herbert	Frank	1920	anglais
3	Azimov	Isaac	1920	anglais
4	Huxley	Aldous	1894	anglais
5	Bradbury	Ray	1920	anglais
6	K. Dick	Philip	1928	anglais
7	Barjavel	René	1911	français
8	Boulle	Pierre	1912	français

Relation : AUTEURS

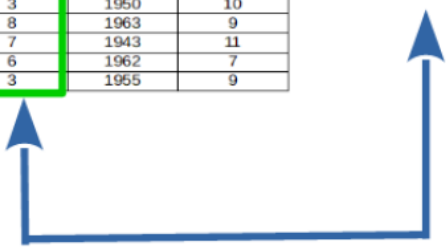


Schéma relationnel

On appelle **schéma relationnel** l'ensemble des relations présentes dans une base de données, c'est-à-dire qu'il faut fournir :

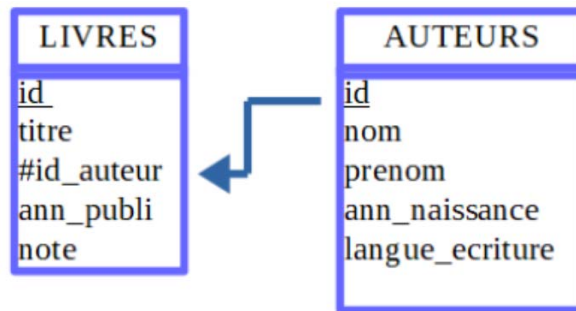
- les **noms** des relations
- la **liste des attributs** et le **type** de leur domaine
- la **clé primaire** (on la souligne)
- les **clés étrangères** (précédées d'un #)

Pour notre exemple cela donne :

LIVRES (id : entiers positifs, **titre** : str, #**id_auteur** : entiers positifs, **ann_publi** : entiers positifs, **note** : entiers positifs)

AUTEURS (id : entiers positifs, **nom** : str, **prenom** : str, **ann_naissance** : entiers positifs, **langue_ecriture** : str)

Ce qui peut aussi se représenter avec ce schéma :



Exercice 1 :

Voici le schéma relationnel approximatif d'une base de données:

<p>Voyageur (idVoyageur, nom, prénom, ville, région) Séjour (idSéjour, idVoyageur, codeLogement, début, fin) Logement (code, nom, capacité, type, lieu) Activité (codeLogement, codeActivité, description)</p>
--

1. Combien y-a-t-il de **relations** ?
2. Citer toutes les **clés primaires**
3. Citer toutes les **clés secondaires** et expliquer leur utilité
4. Réécrire de façon correcte ce **schéma relationnel**

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Exercice 2 :

Voici le schéma relationnel approximatif d'une base de données :

Contact (idContact, nom, prénom, email)
Message (idMessage, contenu, dateEnvoi, idEmetteur, idPrédecesseur)
Envoi (idMessage, idContact)

1. Combien y-a-t-il de **relations** ?
2. Quels sont les **attributs de** la relation **Contact** ?
3. Quel est le **type de** l'attribut **contenu de** la relation **Message** ?
4. Quelle est la **clé primaire** de la relation **Envoi** ?
5. Quelle est la **nature de** l'attribut **IdEmetteur** ? A quoi sert il ?

.....
.....
.....
.....
.....
.....
.....

Établir des schémas relationnels

Exercice 3 :

Un laboratoire souhaite gérer les médicaments qu'il conçoit.

- Un **médicament** est décrit par un **nom**, qui permet de l'identifier. En effet il n'existe pas deux médicaments avec le même nom.
- Un **médicament** comporte une **description courte en français**, ainsi qu'une **description longue en latin**.
- On gère aussi le **conditionnement** du médicament, c'est à dire le nombre de pilules par boîte (qui est un **nombre entier**).

À chaque **médicament** on associe une liste de **contre-indications**, généralement plusieurs, parfois aucune.

Une **contre-indication** comporte un **code unique** qui l'identifie, ainsi qu'une **description**.

Une **contre-indication** est toujours associée à un et **un seul médicament**.

Exemples de données

Voici deux exemples de données :

- Le **Chourix** a pour description courte "« Médicament contre la chute des choux »" et pour description longue "« Vivamus fermentum semper porta. Nunc diam velit, adipiscing ut tristique vitae, sagittis vel odio. Maecenas convallis ullamcorper ultricies. Curabitur ornare. »". Il est conditionné en boîte de 13. Ses contre-indications sont :
 - CI1 : Ne jamais prendre après minuit.
 - CI2 : Ne jamais mettre en contact avec de l'eau.

