

2__La__BDD__relationnelle

September 3, 2020

Base de données : Introduction

Les bases de données relationnelles sont le plus utilisées au monde, elles ont été mises au point en 1970 par Edgar Franck Codd, informaticien britannique (1923-2003). Ces bases de données sont basées sur la théorie mathématique des ensembles.

La notion de relation est au coeur des bases de données relationnelles. Une relation peut être vue comme un tableau à 2 dimensions, composé d'un en-tête et d'un corps. Le corps est lui-même composé de **t-uplets** (lignes) et d'**attributs** (colonnes). Le nombre d'attributs est appelé **degré** de la relation (table)

À noter que l'on emploie aussi le terme “**table**” à la place de “**relation**”.

id	titre	auteur	ann_publi	note
1	1984	Orwell	1949	10
2	Dune	Herbert	1965	8
3	Fondation	Asimov	1951	9
4	Le meilleur des mondes	Huxley	1931	7
5	Fahrenheit 451	Bradbury	1953	7
6	Ubik	K.Dick	1969	9
7	Chroniques martiennes	Bradbury	1950	8
8	La nuit des temps	Barjavel	1968	7
9	Blade Runner	K.Dick	1968	8
10	Les Robots	Asimov	1950	9
11	La Planète des singes	Boulle	1963	8
12	Ravage	Barjavel	1943	8
13	Le Maître du Haut Château	K.Dick	1962	8
14	Le monde des Â	Van Vogt	1945	7
15	La Fin de l'éternité	Asimov	1955	8
16	De la Terre à la Lune	Verne	1865	10

Annotations du schéma :

- Le premier rang (en-tête) est encadré en rouge et pointé par l'étiquette "en-tête".
- Le corps de la table (rang 2 à 16) est encadré en vert et pointé par l'étiquette "corps".
- Le rang 11 est encadré en jaune et pointé par l'étiquette "1 t-uplet".
- Le titre "titre" est pointé par une flèche bleue venant d'en dessous, avec l'étiquette "l'attribut 'titre'".

Le **t-uplet** encadré en jaune sur le schéma ci-dessus contient les éléments suivant : 11, La Planète des singes, Boulle, 1963 et 8.

L'**attribut** "titre" est composé des éléments suivants : 1984, Dune, Fondation, Le meilleur des mondes, Fahrenheit 451, Ubik, Chroniques martiennes, La nuit des temps, Blade Runner, Les Robots, La Planète des singes, Ravage, Le Maître du Haut Château, Le monde des Â, La Fin de l'éternité et De la Terre à la Lune.

Faites la liste des éléments appartenant à l'**attribut** "auteur". Quel est le degré de cet relation ?

0.0.1 Domaine de l'attribut

Pour chaque **attribut** d'une relation, il est nécessaire de définir un **domaine** : Le domaine d'un attribut donné correspond à un ensemble fini ou infini de valeurs admissibles.

Par exemple, le domaine de l'attribut "id" correspond à l'ensemble des entiers (la colonne "id" devra obligatoirement contenir des entiers). Autre exemple, le domaine de l'attribut "titre" correspond à l'ensemble des chaînes de caractères. Dernier exemple, le domaine de l'attribut "note" correspond à l'ensemble des entiers positifs.

selon vous, quel est le **domaine** de l'attribut "auteur" ?

0.0.2 contrainte

Au moment de la création d'une relation, il est nécessaire de renseigner le domaine de chaque attribut. Le SGBD s'assure qu'un élément ajouté à une relation respecte bien le domaine de l'attribut correspondant : si par exemple vous essayez d'ajouter une note non entière (par exemple 8.5), le SGBD signalera cette erreur et n'autorisera pas l'écriture de cette nouvelle donnée.

Autre contrainte très importante dans les bases de données relationnelles, une relation ne peut pas contenir 2 t-uplets identiques. Par exemple, la situation ci-dessous n'est pas autorisée (ici aussi c'est le SGBD qui veille au grain) :

id	titre	auteur	ann_publi	note
1	1984	Orwell	1949	10
2	Dune	Herbert	1965	8
2	Dune	Herbert	1965	8
3	Fondation	Asimov	1951	9

Afin d'être sûr de respecter cette contrainte des t-uplets identiques, on définit la notion de "clef primaire".

Une **clef primaire** est un attribut dont la valeur permet d'identifier de manière unique un **t-uplet** de la relation. Autrement dit, si un attribut est considéré comme clef primaire, on ne doit pas trouver dans toute la relation 2 fois la même valeur pour cet attribut.

Relation LIVRES

id	titre	auteur	ann_publi	note
1	1984	Orwell	1949	10
2	Dune	Herbert	1965	8
3	Fondation	Asimov	1951	9
4	Le meilleur des mondes	Huxley	1931	7
5	Fahrenheit 451	Bradbury	1953	7
6	Ubik	K.Dick	1969	9
7	Chroniques martiennes	Bradbury	1950	8
8	La nuit des temps	Barjavel	1968	7
9	Blade Runner	K.Dick	1968	8
10	Les Robots	Asimov	1950	9
11	La Planète des singes	Boulle	1963	8
12	Ravage	Barjavel	1943	8
13	Le Maître du Haut Château	K.Dick	1962	8
14	Le monde des Â	Van Vogt	1945	7
15	La Fin de l'éternité	Asimov	1955	8
16	De la Terre à la Lune	Verne	1865	10

L'attribut "note" peut-il jouer le rôle de clef primaire ?

L'attribut "ann_publi" peut-il jouer le rôle de clef primaire ?

L'attribut "auteur" peut-il jouer le rôle de clef primaire ?

L'attribut "titre" peut-il jouer le rôle de clef primaire ?

(Voir Exercice 1)

0.0.3 Lien entre relations

Nous avons enrichi la relation livre en ajoutant des informations supplémentaires sur les auteurs, nous obtenons alors :

Relation LIVRES_AUTEURS							
id	titre	nom_auteur	prenom_auteur	date_nai_auteur	langue_ecriture_auteur	ann_publi	note
1	1984	Orwell	George	1903	anglais	1949	10
2	Dune	Herbert	Frank	1920	anglais	1965	8
3	Fondation	Asimov	Isaac	1920	anglais	1951	9
4	Le meilleur des mondes	Huxley	Aldous	1894	anglais	1931	7
5	Fahrenheit 451	Bradbury	Ray	1920	anglais	1953	7
6	Ubik	K.Dick	Philip	1928	anglais	1969	9
7	Chroniques martiennes	Bradbury	Ray	1920	anglais	1950	8
8	La nuit des temps	Barjavel	René	1911	français	1968	7
9	Blade Runner	K.Dick	Philip	1928	anglais	1968	8
10	Les Robots	Asimov	Isaac	1920	anglais	1950	9
11	La Planète des singes	Boulle	Pierre	1912	français	1963	8
12	Ravage	Barjavel	René	1911	français	1943	8
13	Le Maître du Haut Château	K.Dick	Philip	1928	anglais	1962	8
14	Le monde des Â	Van Vogt	Alfred Elton	1912	anglais	1945	7
15	La Fin de l'éternité	Asimov	Isaac	1920	anglais	1955	8
16	De la Terre à la Lune	Verne	Jules	1828	français	1865	10

Nous avons ajouté les attributs (“prenom_auteur”, “date_nai_auteur” et “langue_ecriture_auteur”). Nous avons aussi renommé l'attribut “auteur” en “nom_auteur”.

Comme vous l'avez peut-être remarqué, il y a pas mal d'informations dupliquées, par exemple, on a 16 fois le nom “Asimov”. Cette duplication est-elle indispensable ? Non !

Est-elle souhaitable ? Non plus !

Si nous dupliquons autant de données inutilement c'est que notre structure ne doit pas être la meilleure. Mais alors, comment faire pour avoir aussi des informations sur les auteurs des livres ?

La solution est relativement simple : travailler avec 2 relations au lieu d'une seule et créer un “lien” entre ces 2 relations :

Comme vous l'avez sans doute remarqué, l'attribut “id_auteur” de la relation LIVRES permet de lier un livre à son auteur. L'introduction d'une relation AUTEURS et la mise en place de liens entre cette relation et la relation LIVRES.

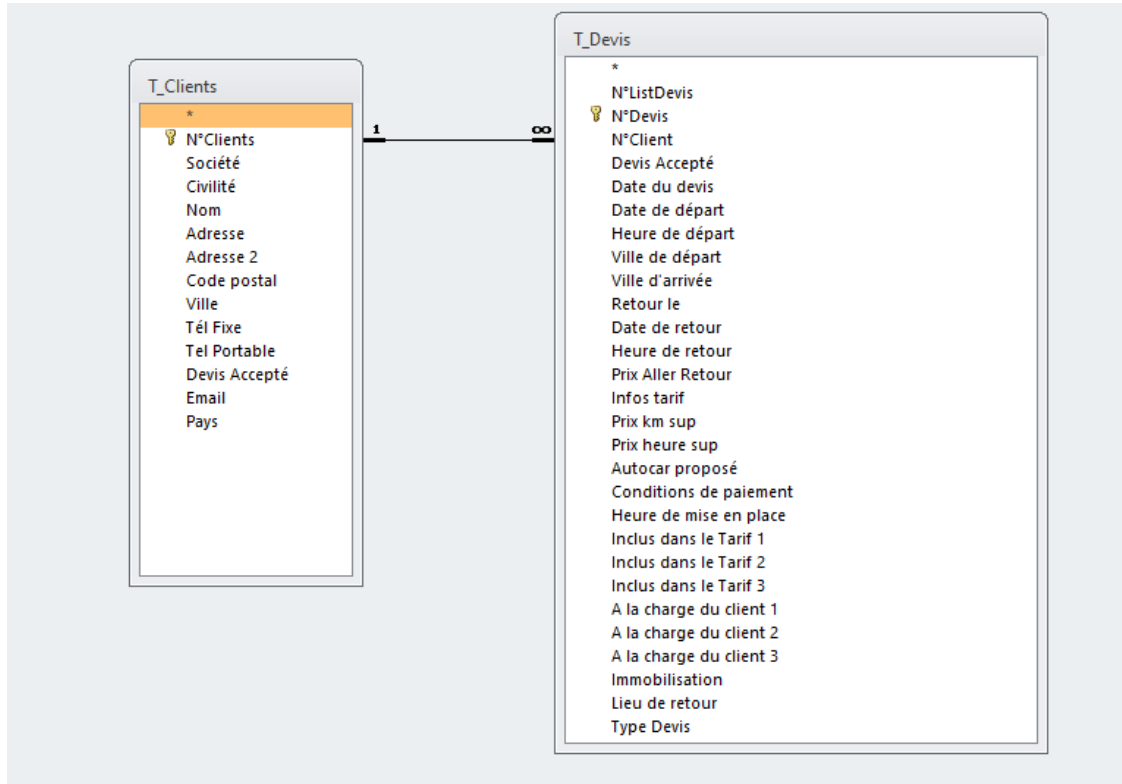
Pour établir un lien entre 2 relations RA et RB, on ajoute à RA un attribut x qui prendra les valeurs des attributs de RB.

Cet attribut x est appelé **clef étrangère** (l'attribut correspond à la **clé primaire** d'une autre table, d'où le nom)

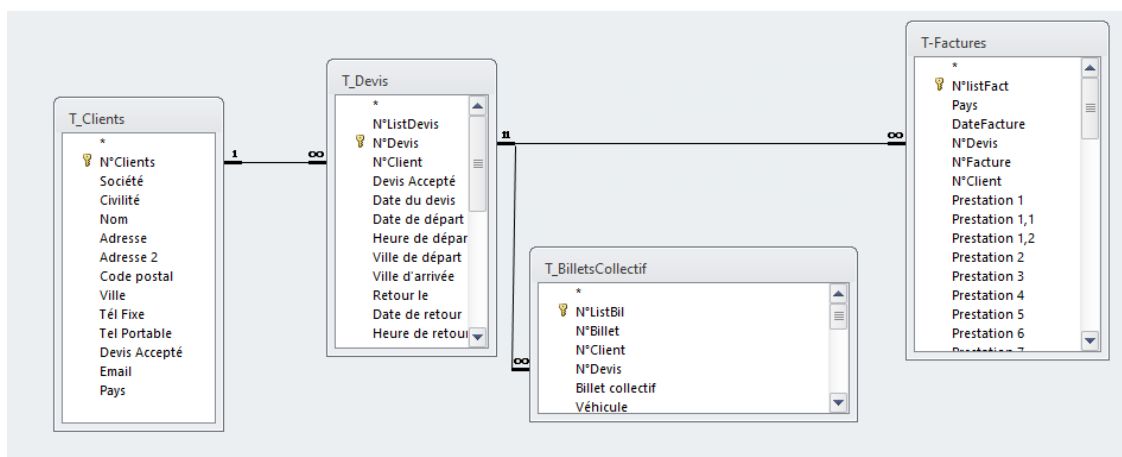
Dans l'exemple ci-dessus, l'attribut “id_auteur” de la relation LIVRES permet bien d'établir un lien entre un livre et son auteur.

0.0.4 Schéma relationnel

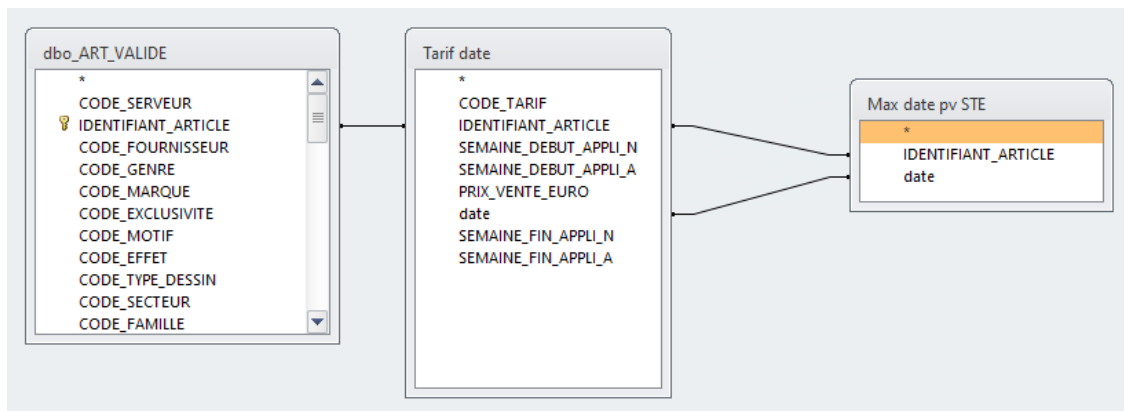
Un **schéma de relationnel** précise le nom de la relation ainsi que la liste des attributs avec leurs domaines. On peut représenter un schéma de relation sous une forme graphique, par exemple à l'aide d'un diagramme



Ci-dessous un autre exemples comprenant quatres relations:



Double jointure.



(Voir exercice 2)

[]: