

## Séance 5 : Bases de données relationnelles I – modèle relationnel

### Remarques préliminaires

Les problèmes rencontrés et les oublis constatés dans ce support sont à signaler à [clot@univ-lyon1.fr](mailto:clot@univ-lyon1.fr)

L'objectif poursuivi dans ce document, et le suivant, est de présenter les concepts du modèle relationnel utilisés dans les bases de données relationnelles, les fonctionnalités que l'on peut attendre d'un SGBD et de familiariser l'étudiant à leur manipulation via Access, le gestionnaire de bases de données de la suite Office de Microsoft.

Il est important de lire complètement le texte d'un item avant de tenter d'atteindre l'objectif visé. Il est fréquent que des indications soient fournies dans l'ensemble de phrases constituant l'item.

### Introduction aux systèmes d'information

Pour quelles raisons la centralisation des données a-t-elle autant de succès dans le monde de l'entreprise ? L'information est un élément critique pour une entreprise, systématiquement pour sa gestion interne, pour la gestion de sa relation client et parfois en raison de la nature du service qu'elle rend. Les gestionnaires de l'information ont longtemps dû faire face à des obstacles récurrents :

- L'éparpillement des données
- La chasse aux incohérences
- Le non respect des contraintes auxquelles sont soumises les données
- L'absence de partage des données entre applications et l'absence de définition de niveaux d'accès
- Les conflits d'accès aux données

Tous ces éléments ont participé à la formalisation du concept de base de données dont une définition est proposée ci-après.

### Base de données et SGBD

Une **base de données** est une collection de données placée sur des supports informatiques et disposant des propriétés suivantes :

- Reposer sur une organisation permettant une **exploitation efficace** des données
- Centraliser toutes les données afin de faciliter leur **partage**
- Définir des **contraintes d'intégrité** traduisant les règles auxquelles sont soumises les données réelles
- Etre utilisable par plusieurs applications **simultanément**
- Contenir les données sur les **entités** et leurs **associations**
- Etre **non-redondante**

La base de données est la structure au sein de laquelle sont organisées les données. L'exploitation de cette structure passe par l'utilisation d'un logiciel appelé « système de gestion de base de données » ou **SGBD**. Plus spécifiquement, un SGBD assure un certain nombre de fonctions dont les plus importantes sont les suivantes :

- Permettre la création et le paramétrage de la base de données
- Permettre les opérations de manipulation telles que l'interrogation et la mise à jour des données
- Assurer le contrôle des règles définies par les contraintes d'intégrité
- Permettre de définir des niveaux de confidentialité pour l'accès aux informations
- Gérer les accès concurrents aux données
- Protéger les données d'événements tels que les pannes

## Le modèle relationnel

Le concept de SGBD s'est essentiellement développé sur sa déclinaison au **modèle relationnel**. D'autres modèles existent, mais c'est celui-ci qui fait l'objet de la majorité d'applications. Le développement de ce type de SGBD a profité des avancées apportées par le paradigme de **client/serveur**, mais le grand succès rencontré par les SGBDs relationnels s'explique par ses caractéristiques propres :

- Le modèle relationnel repose sur des **concepts mathématiques**, chassant toute possibilité d'ambiguïté dans ses implémentations.
- Les concepts définis sont **simples** à appréhender : ils se définissent en termes mathématiques simples et non informatiques autour desquels spécialistes et non-spécialistes peuvent communiquer.
- Le résultat d'une opération peut à son tour être utilisé pour être l'objet d'un traitement, etc. La complexité s'atteint par un ensemble d'opérations simples.
- Comparativement à d'autres solutions de gestion d'informations plus fermées, la solution apportée par un SGBD relationnel permet une **économie** importante.

Un des objectifs du modèle relationnel est de permettre la construction d'une **vue logique** de l'organisation des informations rassemblée par la base sous forme de tables et de liens indiquant la nature des liaisons existant entre ces tables. Cette vue s'appelle le **schéma relationnel**.

L'autre point intéressant du modèle relationnel est qu'il comporte un ensemble d'opérateurs simples et performants pour les opérations d'interrogation de la base. Ces opérateurs fondent l'**algèbre relationnelle**.

## Concepts du modèle relationnel

Nous nous limiterons ici à une présentation des concepts principaux du modèle relationnel en relation avec les usages prévus (compréhension de l'organisation logique des informations et construction de requêtes principalement) dans le cadre d'interaction avec Access dans le support suivant.

### Relation, attributs et tuples

Une base de données est constituée de tables. Une table est constituée de colonnes appelées **attributs** et de lignes appelées **tuples** (ou n-uplets).

La valeur que peut prendre un tuple pour un attribut donné est limité à un ensemble de valeurs pour cet attribut appelé **domaine de l'attribut**. Une table, ou plus formellement une **relation**, est un sous-ensemble du **produit cartésien** des domaines des attributs rassemblés par la relation. En particulier, ceci implique qu'un tuple ne peut être répété dans une table.

Un domaine d'attribut a une signification sémantique particulière. Deux domaines sont dits **compatibles** s'ils sont **sémantiquement comparables**. Le fait que les types de données soient comparables n'est pas suffisant !

Une relation peut être définie en extension (liste des tuples la composant) ou en compréhension (i.e. en caractérisant par un critère d'appartenance à l'ensemble des tuples).

### Clé et domaine primaire

Le fait qu'un tuple soit unique au sein d'une relation implique qu'il est possible d'identifier un attribut ou un ensemble d'attributs permettant d'identifier de façon unique chaque tuple. Une clé primaire est un ensemble d'attributs permettant d'identifier un tuple de façon univoque (propriété d'unicité) et minimal (en ce sens que la suppression d'un attribut de cette clé entraîne la perte de la propriété d'unicité).

Un attribut qui n'est pas une clé primaire mais qui est définie sur un domaine primaire est dit **clé étrangère**. Les clés étrangères permettent de décrire les associations entre tuples de tables différentes.

### Exemple

Prenons l'exemple d'un système d'information rassemblant l'information relative aux consommations de clients d'un établissement de restauration. Les références des clients sont connues ainsi que le détail des plats et boissons prises au cours d'un repas. Les tarifs des consommations sont connus dans le temps, ce qui permet de recalculer les notes de repas alors que les tarifs actuels ont pu évoluer.

Les figures ci-dessous illustrent d'une part l'information éclatée sur différentes tables et le schéma relationnel synthétisant l'organisation de l'information sous forme de tables.

| Repas    |           |      | Tarifs     |      |      |
|----------|-----------|------|------------|------|------|
| Id Repas | Id Client | Date | Id Article | Prix | Date |
|          |           |      |            |      |      |
|          |           |      |            |      |      |

  

| Client    |     |        | Plats   |          |          |
|-----------|-----|--------|---------|----------|----------|
| Id Client | Nom | Prenom | Id Plat | Id Repas | Quantite |
|           |     |        |         |          |          |
|           |     |        |         |          |          |

  

| Boissons   |          |          |
|------------|----------|----------|
| Id Boisson | Id Repas | Quantite |
|            |          |          |
|            |          |          |

Figure 20 - Tables

| Client    | Repas     | Plats      | Tarifs     |
|-----------|-----------|------------|------------|
| Id Client | Id Repas  | Id Plat    | Id Article |
| Nom       | Id Client | Id Repas   | Prix       |
| Prenom    | Date      | Quantite   | Date       |
|           |           | Id Boisson |            |
|           |           | Id Repas   |            |
|           |           | Quantite   |            |

Figure 21 - Schéma relationnel

La table Client rassemble les informations spécifiques au client, la table Repas permet de dater un repas et de faire le lien avec les plats et boissons consommés lors de ce repas. Les tables Plats et Boissons permettent de préciser les quantités consommées et la table Tarifs permet de connaître les prix des consommations avec leur date de mise en vigueur.

- ? Reconstituez l'information sous la forme d'une seule table.
- ? Quel est l'inconvénient majeur de cette organisation de l'information ?
- ? Chacune des tables illustrées sur la Figure 21 met en évidence la clé primaire. Une erreur s'est toutefois glissée. Identifiez-la et proposez une correction.
- ? Comment définiriez-vous le domaine de l'attribut « Id Article » de la table Tarifs ?
- ? Les valeurs pour les attributs « Id Plat » et « Id Boisson » n'étant pas très explicites, il est souhaité d'ajouter un libellé plus lisible pour les plats et les boissons. Proposez une solution que vous jugez mauvaise et une autre meilleure en justifiant ces qualifications.

### Contraintes d'intégrité

Différents types de contraintes existent. C'est le rôle du SGBD de permettre de définir ces contraintes et d'en contrôler le respect. Ces types sont les suivants :

- Les contraintes d'**intégrité de domaine** ont pour objet la restriction des valeurs prises par les tuples pour un attribut aux valeurs contenues dans le domaine associé à cet attribut.
- Les contraintes d'**intégrité de relation** interdisent les doublons dans les valeurs que peuvent prendre les tuples pour la clé primaire. Ce mécanisme vise à assurer la propriété d'unicité des clés primaires.
- Les contraintes de références (ou **référentielle**) permettent de limiter les valeurs pour une clé étrangère aux valeurs déjà utilisées par les clés primaires associées.

- ? Proposez un domaine pour l'attribut « Quantité ».
- ? Quelle précaution particulière faut-il prendre pour la définition des attributs « Id Plat » et Id Boisson » et pourquoi ?
- ? Comment s'exprime la contrainte d'intégrité de relation pour la table « Plat » ?
- ? Comment s'exprime la contrainte référentielle pour les attributs « Id Repas » des tables Repas et Plats ?

### Algèbre relationnelle

Le modèle relationnel propose divers opérateurs : les opérateurs sémantiques (de création et manipulation de domaine) d'une part et les opérateurs ensemblistes et relationnels d'autre part, ces derniers formant l'algèbre relationnelle. Nous nous attacherons essentiellement à ces derniers opérateurs, pour leur rôle important dans l'activité d'interrogation des bases.

### Opérateurs ensemblistes

L'union, l'intersection, la différence et le produit cartésien sont les opérateurs ensemblistes binaires qui s'appliquent à des relations et qui fournissent en résultat une relation :

- L'union de deux relations T1 et T2 dont les attributs sont compatibles deux à deux, notée  $T1 \cup T2$ , est l'ensemble des tuples appartenant à T1 ou à T2 ou aux deux ;
- La différence de deux relations T1 et T2 (dans cet ordre) dont les attributs sont compatibles deux à deux, notée  $T1 - T2$ , est l'ensemble des tuples appartenant à T1 et n'appartenant pas à T2.
- Le produit cartésien de deux relations T1 et T2 est une relation ayant pour attributs tous les attributs de T1 et tous les attributs de T2 et formée de l'ensemble des tuples que l'on peut obtenir par concaténation de chaque tuple de T1 avec chaque tuple de T2.
- ? Considérez deux relations T1 et T2 dont les attributs sont deux à deux compatibles. Comment pouvez-vous écrire l'intersection des relations T1 et T2, i.e. la relation constituée des tuples appartenant à la fois à T1 et à T2, en fonction des autres opérateurs ?

| Boissons      |               |          | Tarifs        |            |              |
|---------------|---------------|----------|---------------|------------|--------------|
| Id Boisson    | Id Repas      | Quantité | Id Article    | Prix       | Date         |
| <b>B00001</b> | <b>R00001</b> | <b>1</b> | <b>B00001</b> | <b>4.6</b> | <b>08/01</b> |
| <b>B00002</b> | <b>R00001</b> | <b>2</b> | <b>B00002</b> | <b>5.2</b> | <b>08/01</b> |

Figure 22 - Données illustratives

- ? Considérez les tables de la Figure 22 et déterminez le produit cartésien de ces deux tables.

### Opérateurs relationnels

Les opérateurs relationnels de sélection et de projection se définissent comme suit :

- La projection d'une relation T d'attribut  $C1, C2 \dots Cn$  sur les attributs  $Ci_1, Ci_2 \dots Ci_p$  se définit comme la relation sur les attributs  $Ci_1, Ci_2 \dots Ci_p$  dont les tuples sont obtenus par élimination dans les tuples de T des valeurs des attributs non sélectionnés pour la projection et après suppression des éventuels doublons.
- La sélection d'une relation T selon un certain critère est la relation formée des tuples de T satisfaisant le critère de sélection.
- ? Quelle est la projection de la table Boisson sur l'attribut « Id Boisson » ? Et sur l'attribut « Id Repas » ?
- ? Quelle est la projection sur les attributs « Id Boisson » et « Prix » de sélection du produit cartésien des deux relations pour laquelle les valeurs pour les attributs « Id Boisson » et « Id Article » sont égales ?

### Opérateurs additionnels

Les opérateurs ensemblistes et relationnels sont suffisants pour exprimer toutes les opérations de l'algèbre relationnelle. Cependant, certaines combinaisons reviennent de façon récurrente dans les interrogations et ont conduit à la définition d'opérateurs additionnels :

- La **jointure** de deux relations T1 et T2 selon un critère C est l'ensemble des tuples obtenus par produit cartésien de T1 par T2 et satisfaisant le critère C.
- Certains critères classiques ont donné lieu à des déclinaisons de la jointure :
  - L'**équijointure** de T1 et T2 sur les attributs compatibles A1 de T1 et A2 de T2 tels que  $A1=A2$  ;
  - La **jointure naturelle** de T1 et T2 est l'équijointure de T1 et T2 sur tous les attributs portant le même nom dans T1 et dans T2, suivie d'une projection permettant de supprimer la répétition de ces attributs.
- La **jointure externe** de deux relations T1 et T2 selon le critère C est la jointure de T1 et T2 selon C à laquelle s'ajoutent les tuples de T1 et T2 n'apparaissant pas dans cette première jointure et complétés par des valeurs vides.
- La **division** de la relation T1 d'attributs  $C_1 C_2 \dots C_p C_{p+1} \dots C_n$  par la relation T2 d'attributs  $C_{p+1} \dots C_n$  est la relation d'attribut  $C_1 C_2 \dots C_p$  formée des tuples qui étaient associés, par concaténation, à chacun des tuples de T2.

- ? Comment utiliseriez-vous les opérateurs précédents pour reconstruire l'information sous la forme d'une seule table.
- ? Comment composeriez-vous les opérations pour connaître les clients qui ont goûté à tous les plats ?
- ? Comment déterminer les plats qui n'ont jamais été commandés ?
- ? Comment identifier les plats qui ont été associés à chacune des boissons.
- ? Comment déterminer la liste des repas au cours desquels seules des boissons ont été consommées ?