

Initiation au langage SQL

José Arthur OUEDRAOGO

École Supérieure d'Informatique (ESI)

josearthur.oued@outlook.com

January 11, 2022



Plan

- 1 Aperçu du cours
 - Prérequis
 - Objectifs du cours
- 2 Définition d'une base de données - SGBD
 - Définition d'une base de données
 - Systèmes de Gestion de Bases de Données (SGBD)
- 3 Découvrir le langage SQL
 - Historique du langage SQL
 - Structure d'une instance : base, schémas et objets
 - Principaux SGBD (Systèmes de Gestion de Base de Données)
- 4 Concevoir une base de données
 - Modèle Conceptuel de données : tables et relations
 - Passage du Modèle Conceptuel de donnée (MCD) au Modèle Logique de donnée (MLD)

- 1 Aperçu du cours
 - Prérequis
 - Objectifs du cours
- 2 Définition d'une base de données - SGBD
- 3 Découvrir le langage SQL
- 4 Concevoir une base de données

Prérequis

- Avant de commencer ce module, l'apprenant doit avoir les bases suivantes: compétences basiques dans les TIC, les bases mathématiques comme la théorie des ensembles et l'introduction à la programmation.

Objectifs du cours

À la fin de ce cours, l'étudiant devrait être en mesure de :

- Décrire, définir et appliquer les composants majeurs du modèle de base de données relationnelle pour la conception de la base de données;
- Apprendre et appliquer le langage SQL (Structured Query Language) pour la définition et la manipulation des bases de données;
- Découvrir les commandes de base du langage SQL pour réaliser des requêtes d'interrogation et ainsi obtenir des informations sur une relation d'une base de données.

- 1 Aperçu du cours
- 2 Définition d'une base de données - SGBD
 - Définition d'une base de données
 - Systèmes de Gestion de Bases de Données (SGBD)
- 3 Découvrir le langage SQL
- 4 Concevoir une base de données

Base de données

- Une base de données est un ensemble structuré d'informations stockées de manière permanente. De ce fait, l'information contenue dans la base représente la banque de données. Ainsi, la base de données est le contenant et la banque de données est le contenu. Une base de données permet de rassembler et de centraliser l'information nécessaire aux diverses applications pour une meilleure distribution (Mathieu, 2000).
- Logiquement, une base de données est composée d'un ensemble d'informations concernant un thème spécifique. Cet ensemble doit être complet, non redondant et structuré. En d'autres termes, toutes les informations sur le sujet donné doivent être disponibles, chaque information doit être unique et structurée.

SGBD

Un SGBD est un logiciel qui prend en charge la structuration, le stockage, la mise à jour et la maintenance d'une base de données. Il est l'unique interface entre les informaticiens et les données (définition des schémas, programmation des applications), ainsi qu'entre les utilisateurs et les données (consultation et mise à jour).

- 1 Aperçu du cours
- 2 Définition d'une base de données - SGBD
- 3 Découvrir le langage SQL**
 - Historique du langage SQL
 - Structure d'une instance : base, schémas et objets
 - Principaux SGBD (Systèmes de Gestion de Base de Données)
- 4 Concevoir une base de données

Historique, versions et normalisation du langage SQL

- Les **BD** [Base de Données] sont nés vers la fin des années 1960 pour combler les limites des systèmes de fichiers. Les BD relationnelles, issues de la recherche de Codd, sont celles qui ont connu le plus grand essor et qui reste encore aujourd'hui les plus utilisées.
- Le langage **SQL** [Structured Query Language] est une couche technologique, idéalement indépendante des implémentations des SGBDR [Système de Gestion de Bases de Données Relationnelles], qui permet de créer et manipuler des BD relationnelles.
- **SQL** a été adopté comme recommandation par l'Institut de normalisation américaine (ANSI) en **1986**, puis comme norme internationale par l'ISO en 1987 sous le nom de ISO/CEI 9075 - Technologies de l'information - Langages de base de données - SQL.

Instance et schéma (1/2)

- Une **instance** est l'information recueillie dans une base de données à un moment précis. C'est l'état actuel ou l'occurrence d'une base de données qui est gardé à un moment donnée. Chaque fois que les données sont insérées ou supprimées d'une base de données, l'état d'une base de données est modifié, ce qui explique pourquoi l'instance d'une base de données change très souvent.

Instance et schéma (2/2)

- Le schéma d'une base de données est spécifié dans le SGBD lorsqu'une nouvelle base de données est définie, à ce moment la base de données correspondante est vide, d'où une instance vide. L'état de départ d'une base de données est acquis lorsque la base de données est chargée pour la première fois avec les données initiales. À partir de ce moment, chaque fois que les données sont mises à jour, nous obtenons une nouvelle instance de base de données. À tout moment, il y a un état actuel associé à une base de données. Le SGBD est partiellement responsable de la confirmation de l'instance valide d'une base de données dans laquelle l'instance assure la structure et les contraintes spécifiées dans le schéma.

Principaux SGBD(1/2)

Il y a différentes façons de classer les SGBD, en voici une reposant sur des aspects pratiques, avec des avantages et des inconvénients:

- Le **SGBD relationnel**, le plus fréquent actuellement, repose sur la théorie mathématique des ensembles. Ce système fiable et robuste est dominé depuis trente ans par de grands acteurs, comme Oracle, Microsoft et IBM. Il garantit de bonnes performances en ce qui concerne le stockage, les accès et la sécurité des données, mais avec l'avènement du Big Data, il s'avère rigide et limité, ce qui remet en cause sa prédominance.
- Le **système NoSQL**, plus flexible, il ne demande pas de définir chaque donnée pour chaque entité, ce qui en fait une option intéressante pour gérer des bases de données importante, avec beaucoup de données éparses et dont la structure est susceptible d'évoluer au fil du temps. Il regroupe quatre catégories de fonctions principales : clé-valeurs, documents, colonnes et graphiques.

Principaux SGBD (2/2)

- Le Système de gestion de BD in-Memory, dont l'intérêt principale est de se passer de disque dur et de stocker les données dans sa mémoire. Il est beaucoup plus rapide pour traiter les opérations, mais sa capacité et ses fonctionnalités sont limitées.
- D'autres systèmes, comme le SGBD pour les données XML, peu avantageux par rapport aux autres, le système avec une base de données en colonnes, optimisé et rapide en lecture, mais pas en écriture, les SGBD orientés objets, populaires dans les années 1990, ou enfin, les systèmes hiérarchiques ou les systèmes de réseaux, antérieurs aux bases de données relationnelles.

Modèle Conceptuel de données(1/2)

Un modèle conceptuel de données (MCD) pour une base de données permet d'identifier les principales entités à représenter, leurs relations et leurs attributs, et d'analyser la structure conceptuelle du système d'information. Contrairement à un modèle logique de données (MLD) ou à un modèle physique de données (MPD), un MCD est plus abstrait.

Modèle Conceptuel de données(1/6)

Les MCD offrent les avantages suivants :

- Un diagramme de relation entre entités (ERD) est une représentation graphique de la relation entre les entités de données.
- Valider la conception des données.
- Créez un modèle logique de données (MLD), un modèle physique de données (MDP) ou un modèle orienté objet (OOM), conformément à la norme UML.

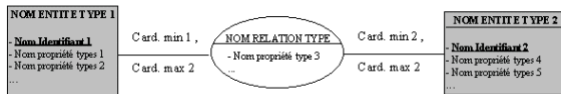
Modèle Conceptuel de données(2/6)

- Le MCD est l'élément le plus connu de MERISE et certainement le plus utile. Il permet d'établir une représentation claire des données du SI et définit les dépendances fonctionnelles de ces données entre elles. Les éléments utilisés pour la formalisation d'un MCD sont les suivants :

Modèle Conceptuel de données(3/6)

Entité Type	Définition d'entités (objets physiques ou abstraits) ayant des caractéristiques comparables.
Relation Type	Définition d'une Association liant plusieurs Entités Types. Signification d'un lien entre deux ou plusieurs types d'objets.
Propriété Type	Définition d'une caractéristique d'un objet ou d'une association. Une propriété Type est elle-même caractérisé par un type (Chiffre ou Texte ...) et une longueur. L'ensemble des propriétés types du MCD compose le <u>dictionnaire des données</u> .
Identifiant	Propriété Type ou concaténation de Propriétés Types permettant de distinguer une entité parmi toute les autres dans une Entité Type.
Cardinalité minimum	Nombre minimum de fois où une entité est concernée par l'association. 0 indique que les entités ne sont pas obligatoirement concernés par l'association.
Cardinalité maximum	Nombre maximum de fois où une entité est concernée par l'association. N signifie plusieurs fois sans préciser de nombre. Ce nombre ne peut être égal à 0.

Représentation



Modèle Conceptuel de données(5/6)

Règles de gestion

Hormis les règles de gestion définies au sein des contraintes d'intégrités représentées par le Schéma, on distingue :

<u>Contraintes statiques</u>	Forme ou Ensemble des valeurs d'une propriété. <u>Exemple</u> : Le poids d'un pilote doit être compris dans la fourchette poids minimum, poids maximum autorisé par le MODELE DE PARAPENTE du PARAPENTE qu'il loue au club.
<u>Contraintes dynamiques</u>	Indiquent les restrictions lors de l'évolution du <u>SI</u> . <u>Exemple</u> : Le niveau d'un pilote ne peut baisser.

...

Modèle Conceptuel de données(6/6)

Règles à suivre pour l'établissement d'un MCD

Normalisation

<u>1^{ère} forme Normale</u>	<p>Chaque entité doit disposer d'un identifiant qui la caractérise de manière unique.</p> <p>Le numéro de licence est unique pour chaque PILOTE</p>
<u>2^{ème} forme normale</u>	<p>Les propriétés d'une entité ne doivent dépendre que de l'identifiant de l'entité et non d'une partie de cet identifiant. Un identifiant peut être composé de la concaténation de plusieurs propriétés.</p>
<u>3^{ème} forme Normale</u>	<p>Les propriétés d'une entité doivent dépendre de l'identifiant de l'entité de manière directe.</p> <p><u>Exemple</u> : Si l'on rajoutait dans l'entité PILOTE une propriété "Description du Niveau" cette normalisation ne serait pas respectée car la Description du Niveau est directement dépendante du Niveau et non du "numéro de licence".</p>
<u>Forme Normale de BOYCE-CODD</u>	<p>Pour les identifiants composés de plusieurs propriétés, ces dernières ne doivent pas être dépendantes d'une autre propriété de l'entité.</p>
<u>Normalisation des relations</u>	<p>Les propriétés des relations doivent dépendre de tous les identifiants des entités associées.</p>
<u>Décomposition des relations</u>	<p>Les relations dont le nombre d'entités associé est trop important (supérieur à 3) doivent être décomposées en plusieurs relations.</p> <p>Cette décomposition ne peut se faire qu'à la condition d'avoir une cardinalité minimum égale à 1.</p>

Passage du MCD au MLD

Un schéma entité- association peut être traduit mécaniquement en un schéma relationnel. (Certains logiciels, WinDesign, BD Designer ... le font automatiquement). Nous allons voir ici comment pratiquer "à la main" pour passer du Modèle Conceptuel des Données entité-association au modèle relationnel.