

Introduction aux bases de données

Introduction

Programme du cours

- Introduction aux BD et aux SGBD
- Le modèle relationnel
- Le langage de requête SQL
- La conception d'une BD relationnelle

Programme du cours

- **Introduction aux BD et aux SGBD**
- Le modèle relationnel
- Le langage de requête SQL
- La conception d'une BD relationnelle

Introduction : pourquoi utiliser une BD

Nécessité d'une bonne organisation des services d'une entreprise, ou institution et surtout la digitalisation des services (applications et logiciels)

Exemple

- Une entreprise conserve un volume élevé d'information: (noms, adresses, salaire, adresse des fournisseurs, quantités, prix des items, bilan financier, etc.)
- Ces informations se retrouvent dans différents systèmes (système de gestion des stocks, système de facturation, système de préparation de paie, RH , etc.)

Introduction : pourquoi utiliser une BD

Exemple (suite)

La recherche d'une information nécessite de :

- déterminer le système à consulter

Certaines informations sont conservées dans plusieurs systèmes

- Duplication de données
- Explosion du volume de fichiers

Inconvénients des systèmes de traitement de fichiers

1. Redondance et inconsistance des données

- Informations identiques répliquées dans plusieurs fichiers

Exemple : coordonnées d'un employé :

 dans le fichier du système de paie

 dans le fichier de gestion du personnel

- Accroissement inutile : de la taille des fichiers, des temps d'accès
- Risque d'inconsistance des données si le changement d'adresse ne s'effectue pas dans les deux fichiers

Inconvénients des systèmes de traitement de fichiers

2. Difficulté d'accès aux données

- Il faut un programme spécifique pour toute nouvelle demande d'information.

Exemple : Estimation du CA pour l'augmentation de 10% sur le prix des produits

- Le temps d'accès à une requête non prévue peut être très long.

Inconvénients des systèmes de traitement de fichiers

3. Isolement des données

- Les données sont stockées sous différents formats

Exemple : Numéro d'un employé

type caractère dans un fichier

type entier dans un autre fichier

- Grande difficulté d'écrire un programme d'accès général à toute l'information

Inconvénients des systèmes de traitement de fichiers

4. Multiplicité des mises à jour

- Les traitements concurrents peuvent générer des erreurs

Exemple : Mises à jour du système de stock d'un produit en même temps

Stock initial 500 unités

O1: ajout de 300 unités

O2: retrait de 600 unités

si O1 avant O2 : 500 unités → 800 unités → 200 unités → OK

si O2 avant O1 : 400 unités → -100 unités → 200 unités → ERREUR!!!

- Nécessite un programme pour gérer les mises à jour

Inconvénients des systèmes de traitement de fichiers

5. Sécurité

- La sécurité des données et les accès non-autorisés ne sont pas garanties.
Exemple : Le personnel ne devrait pas avoir accès au programme de paie.

6. Intégrité des données

- difficulté d'imposer des contraintes
Exemple : le stock d'un produit ne doit jamais être inférieur à 0.

Solution : une base de données

Solution : une base de données commune, entièrement centralisée

Objectif : avoir ...

- Un seul exemplaire de chaque élément de données
 - Tous les utilisateurs ont accès aux données en ne communiquant qu'avec la base (sans intermédiaire).
 - Mesures de protection pour l'information confidentielle
 - La complexité du stockage ne doit pas être apparente à l'utilisateur.
- **Les SGBD garantissent toutes ces fonctionnalités**

Systeme de gestion des BD : SGBD

SGBD

un ensemble de programmes permettant à des utilisateurs de créer et d'utiliser les BDs. Les fonctions supportées sont:

- la définition d'une base de données (spécification des types de données, des structures et des contraintes)
- la construction d'une base de données (stockage des données proprement dites)
- la manipulation des données (ajouter, supprimer, retrouver des données).
- le partage d'une base de données (par les utilisateurs et les programmes)

On peut considérer un SGBD comme un interpréteur d'un langage de programmation de haut niveau qui permet à l'utilisateur de décrire précisément ce qu'il veut obtenir et non comment l'obtenir : c'est-à-dire formuler une assertion et non décrire une procédure

Les sept fonctions d'un SGBD

1. Définition des données
2. Manipulation des données
3. Persistance des données
4. Intégrité des données
5. Sécurité de fonctionnement
6. Accès concurrents
7. Confidentialité des données

Fonctions d'un SGBD

1. Permettre la définition des données :

Un Langage de Définition de Données (LDD) permet de décrire :

- Des **objets** (ex : *étudiants, formations, produits*)
- Des **attributs** sur les objets (*nom, adresse, niveau, durée, prix*)
- Des **liens** entre objets (*étudiant inscrit à une formation*)
- Des **contraintes** sur objets, attributs et liens (*un étudiant n'est inscrit qu'à une formation, numerus clausus*)

Schéma d'une BD : description d'une BD à l'aide du LDD

Sous-schéma : partie de la BD visible par un programme/utilisateur

Fonctions d'un SGBD

2. Permettre la manipulation des données

- Un Langage de Manipulation de Données (LMD) permet d'ajouter, *rechercher, supprimer, modifier* des données
- Il existe d'autres façons de manipuler les données
 - ✓ interface graphique orientée utilisateur final
 - ✓ interface orientée programmeur d'applications
(*SQL immergé dans un langage de programmation, API*)

Le langage SQL comporte des commandes permettant de **définir** et de **manipuler** des données relationnelles

Fonctions d'un SGBD

3. Persistance des données

- Durée de vie des données supérieure à celle des programmes qui les ont générées
 - Stockage des données en mémoire secondaire (disque externe)

- Volume des données très supérieur à la capacité de la mémoire
 - Gestion de cache pour un accès rapide aux données
 - Techniques d'optimisation de requêtes

Fonctions d'un SGBD

4. Contrôler l'intégrité des données

Les données enregistrées doivent vérifier certaines propriétés appelées **contraintes d'intégrité** (CI)

- ✓ CI exprimées dans le schéma (définition) de la BD
 - ➔ le SGBD doit les préserver durant la vie de la BD
- ✓ CI plus complexes
 - ➔ à coder dans des programmes d'application

Fonctions d'un SGBD

5. Assurer la sécurité de fonctionnement

- Mécanisme de **transaction**

- Séquence d'opérations faisant passer une BD d'un état cohérent à un nouvel état cohérent
- Une transaction est exécutée complètement avec succès ou est annulée.

- Remettre rapidement une BD dans un état opérationnel après un incident hardware/software

1. **Journalisation** des opérations réalisées sur la BD
(dans le *journal* ou *LOG*)
2. **Ré-exécution** automatique en cas d'incident

Fonctions d'un SGBD

6. Gérer les accès concurrents

- Permettre des accès simultanés par plusieurs utilisateurs
- Gérer les conflits d'accès
 - autorisation d'accès multiples en consultation → (lecture)
 - verrouillage en cas d'accès en modification → (écriture)

Fonctions d'un SGBD

7. Assurer la confidentialité

Mise en commun des données mais restriction des accès (cacher certaines informations à certains utilisateurs)

- système d'authentification (compte & mots de passe)
- privilèges d'accès
- utilisation de sous-schémas (visibilité partielle)

Terminologie BD

- Données
- Modèle de données
- Schéma et Instance d'une BD
- Contraintes d'intégrité
- Métabase

Données

- **Une structure**
 - Simple: prix, nom, date
 - Complexe: personne, document, image
- **Une sémantique**
 - L'état du stock courant d'un produit
- **Un propriétaire**
 - Responsable de la création d'une donnée
 - Définit les règles pour son identification et son intégrité:
 - « Le stock ne peut pas être négatif »
 - Accorde des droits d'utilisation

Schéma et instance d'une BD

- **Schéma (intention) :**
 - Description des données de la base, conformément à un modèle
 - Schéma relationnel, ...
 - Statique en général
- **Instance (extension) :**
 - Collection de données de la base écrite selon un certain modèle
 - Instance du schéma
 - Dynamique

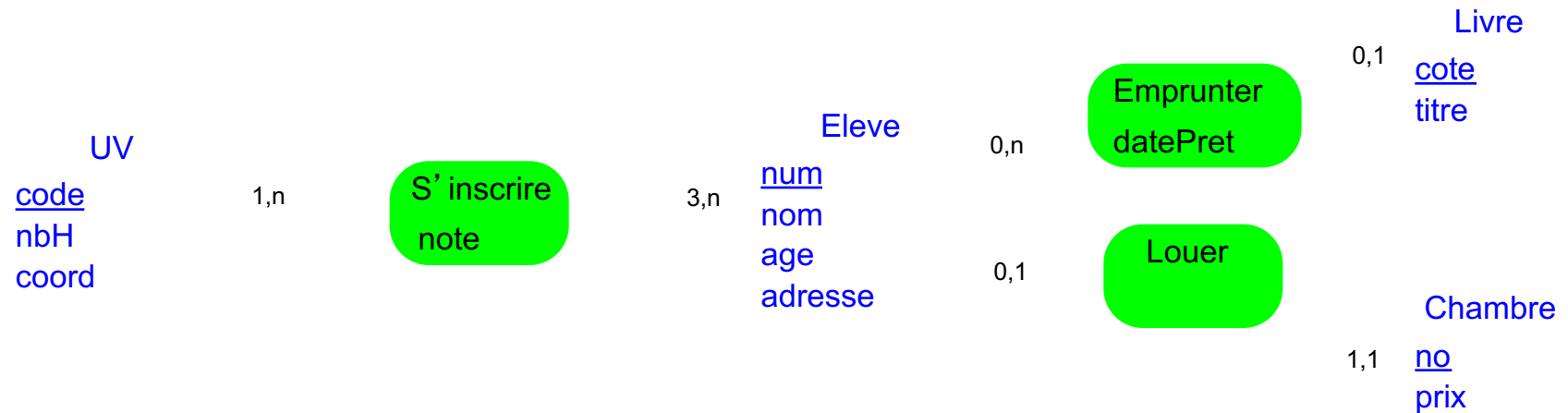
CI et Métabase

- **Contrainte d'intégrité (CI):**
 - Règle spécifiée sur les données, pour définir un état cohérent de la base
 - Le salaire d'un employé doit être supérieur au SMIC
- **Métabase (dictionnaire de données - DD):**
 - Collection des données qui décrivent la BD
 - « Valeur du schéma »

Les modèles de données pour les BD

- **Modélisation des données**
 - Modèle entité/association [Chen 76]
 - Modèle relationnel

Modèle Entité/Association



Modèle Entité/Association

Avantages

- Sémantique riche
- Extension aux concepts objets (héritage, ...)
- Aspect visuel

⇒ Modèle de conception de BD

Inconvénients

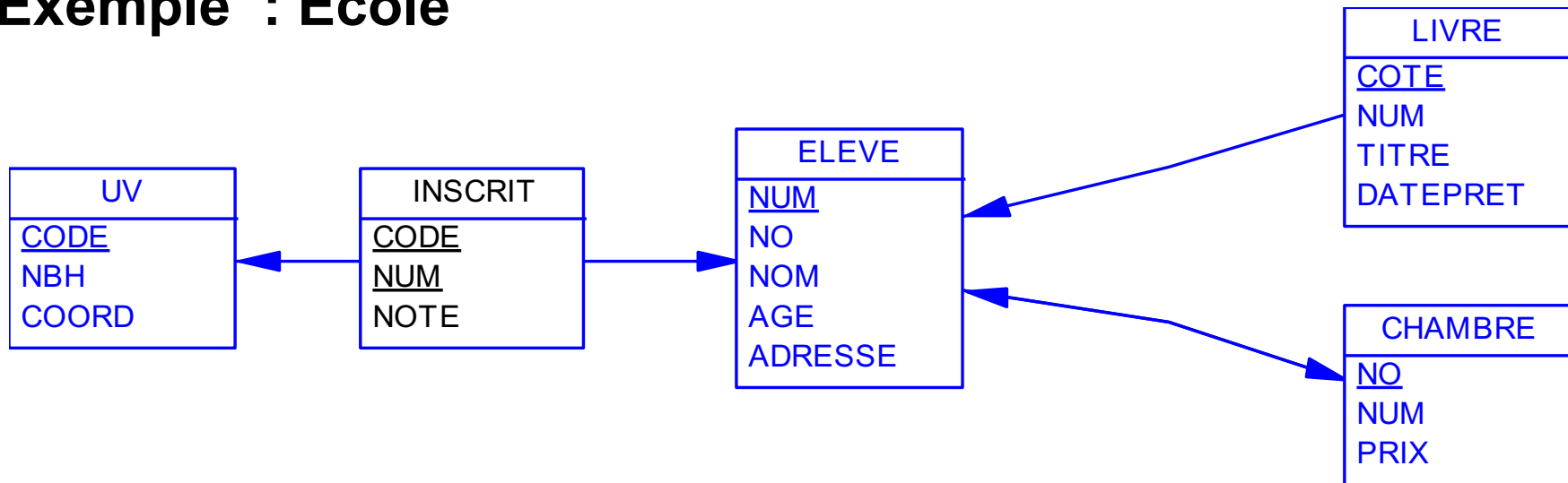
- Uniquement un modèle de description de données
- Pas de langage de manipulation associé
- Pas de SGBD E/A

⇒ Pas un modèle d'implantation de BD

Modèle relationnel

- Schéma BD : Structure de relation
- BD : Ensemble d'enregistrements reliés par des **valeurs**
- Langage de manipulation : Ensembliste , Déclaratif

Exemple : Ecole

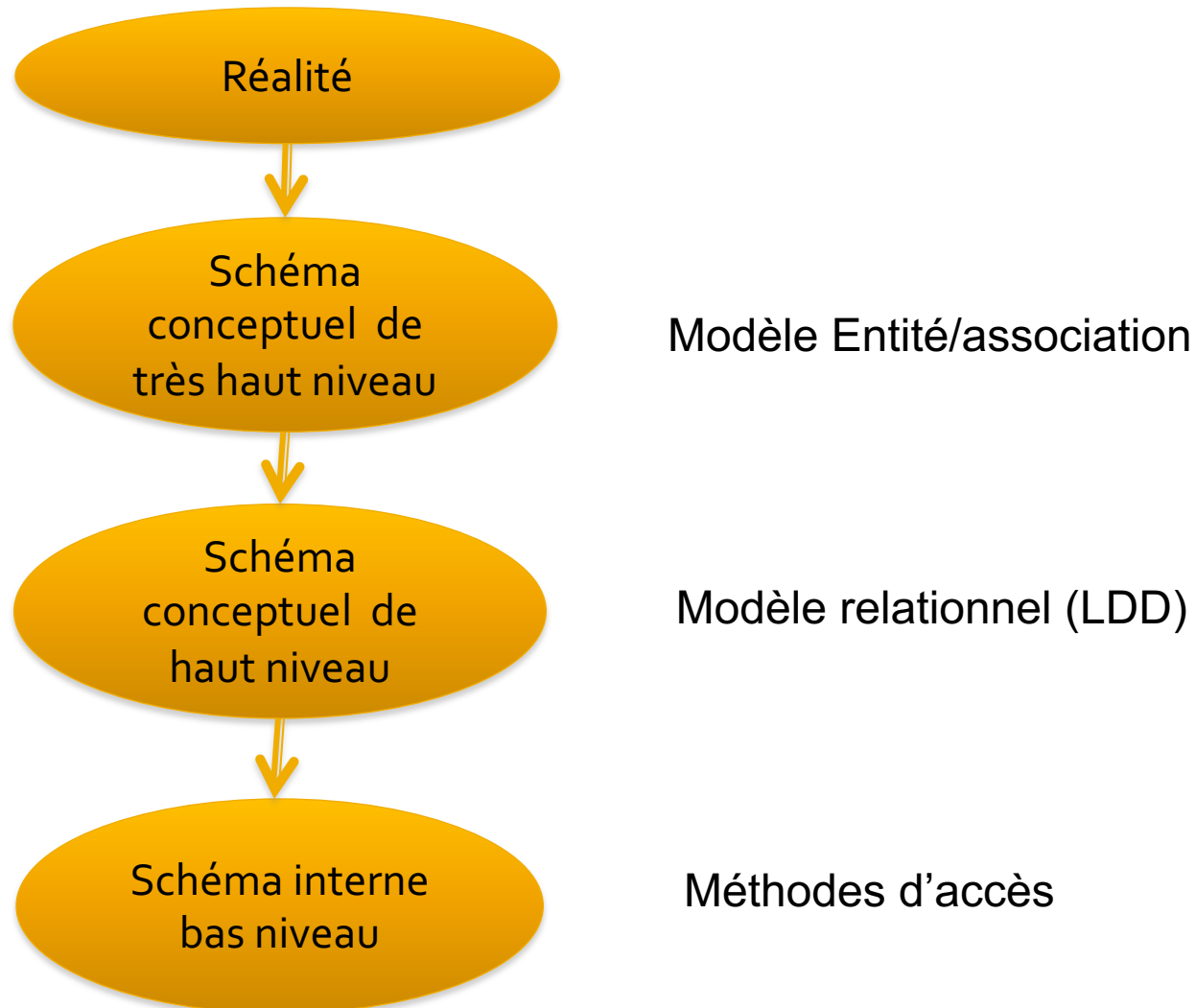


Modèle relationnel (suite)

- **Avantages**

- Indépendance logique/physique
- Langage de manipulation simple
- Basé sur une théorie mathématique solide
- Standard

Conception d'une BD



Architecture d'un SGBD

