

TP 4 — Conception de bases de données

De nombreux exemples et exercices de ce TD sont tirés du livre *Bases de données, Informatique BUT 1^{re} année et L1*, Stéphane DEVISMES, Anaïs DURAND, Pascal LAFOURCADE, Dunod, 2024.

Dans ce TP, nous allons étudier des outils utiles à la conception de bases de données. Avant de créer la base, puis ses tables avec leurs attributs et les relations entre les éléments, il y a une phase de préparation qui peut être assez longue.

Elle consiste à analyser les données à modéliser, afin d'identifier les données importantes et la nature des liens entre elles. C'est cette analyse qui va permettre d'identifier quelles vont être les tables, quels seront les attributs (et donc leur type), les clés primaires, les clés étrangères, les cardinalités, etc.

Il existe de nombreux outils permettant d'aider à la conception d'une base de données. Dans cette ressource introductive, nous allons nous concentrer sur le *modèle conceptuel des données* : c'est une représentation, souvent donnée sous la forme graphique, qui aide à anticiper comment les tables sont agencées entre elles. Le diagramme obtenu à la fin, qui contient les informations nécessaires à la création des tables, s'appelle le *diagramme entités-relations*.

Ce TP se déroule en deux parties :

- 1) La première partie *se fait sans ordinateur* et contient des exercices d'introduction pour se familiariser avec les notions de modèle conceptuel de données et de diagramme entités-relations.
- 2) La deuxième partie consiste à créer une base de données en se basant sur un diagramme que vous aurez préalablement élaboré.

1 Modélisation de données

Le *modèle conceptuel des données* (MCD) est un diagramme représentant un ensemble de données de façon structurée. Il permet de recenser et d'organiser les données manipulées et les relations entre elles.

1.1 Modèle conceptuel des données

Définition. Une *entité* est un groupement logique de données, décrites par un ensemble d'informations élémentaires appelées *attributs*.

Étudiant-es
<u>Numéro d'étudiant-e</u> 
Nom
Prénom
Téléphone
Âge
Année
Groupe

Lorsque nous avons modélisé le BUT informatique, les entités étaient les étudiant-es, les enseignant-es et les enseignants. En général, une entité se traduit par une (ou plusieurs) tables dans la base de données. Il n'y a pas une unique manière de choisir les entités : c'est un choix de conception qui aura des conséquences sur la façon dont la table sera structurée.

Chaque entité doit être munie d'un *identifiant* qui permet distinguer ses éléments. Un identifiant est un attribut ou *un ensemble de plusieurs* attributs, qui formeront ensuite la clé primaire de la table. Par convention, les identifiants sont soulignés.

On représente les entités par des rectangles.

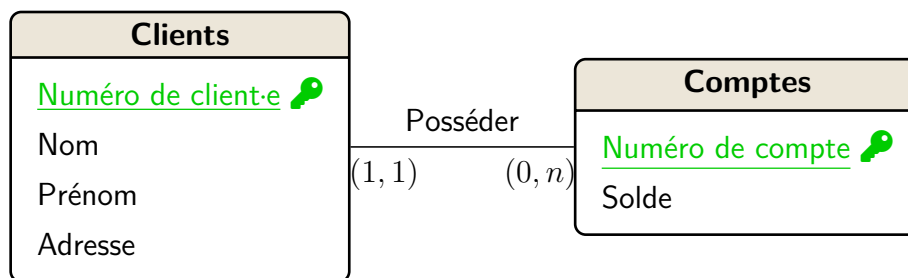
Exercice 1. En se basant sur l'exemple fourni pour l'entité **etudiants**, construire le diagramme des entités pour modéliser le BUT informatique.

Définition. Les entités sont reliées entre elles par des relations appelées *associations*. Les associations sont général désignées par des verbes à l'infinitif et peuvent éventuellement être *orientées*.

Si E1 et E2 sont deux entités reliées par une association, on munit l'association de *cardinalités*. Les cardinalités sont *orientées* (de E1 vers E2 ou de E2 vers E1. Chaque cardinalité est un couple (a, b) de nombres entiers positifs, où :

- a est le nombre minimal de liens entre les deux entités ;
- b est le nombre maximal de liens entre les deux entités.

Exercice 2. Voici un extrait de modèle conceptuel de données issu d'une base de données bancaire :



- 1) Que signifient les cardinalités $(0, n)$ et $(1, 1)$?
- 2) Modifier ces cardinalités pour accepter les comptes joints.
- 3) Modifier les cardinalités pour retirer les client-es n'ayant pas de compte bancaire.

Exercice 3. 1) Reprendre le diagramme des entités élaboré exercice 1 et le compléter avec les relations, munies de leurs cardinalités.

- 2) Préciser la cardinalité de l'association entre les entités **enseignements** et **competences**.

Exercice 4. On souhaite modéliser les données d'une maison de disques. Le cahier des charges indique que l'on souhaite enregistrer des informations relatives aux artistes (nom, prénom, éventuel nom de scène), aux chansons (titre, durée) et aux albums (titre, date de sortie). Chaque chason peut avoir été composée par au minimum zéro artiste (si les auteurs ou autrices sont anonymes), éventuellement plusieurs. Une chanson peut n'appartenir à aucun album (c'est un *single*) ou à un album. Dans ce dernier cas, le titre de l'album est le titre de la première chanson. Proposer un modèle conceptuel de données pour modéliser cette situation.

1.2 Du modèle conceptuel des données au diagramme entités-relations

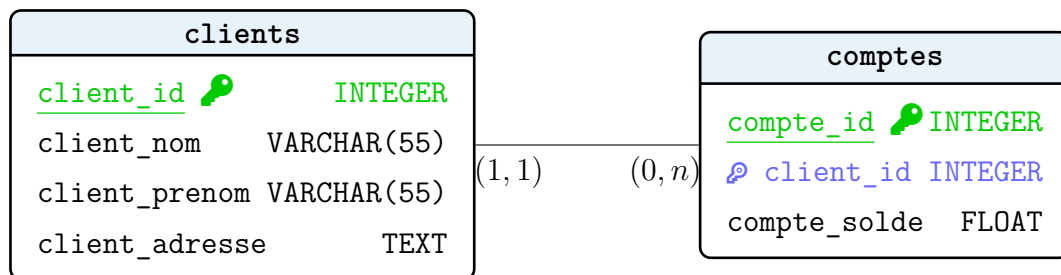
Dans le MCD, on ne manipule aucun vocabulaire lié à SQL. Il s'agit seulement d'organiser les données et de préparer la construction de la base de données. Les identifiants dans les entités deviendront les clés primaires, les cardinalités indiqueront les clés étrangères et les éventuelles tables de relations.

Après avoir conçu le MCD, on construit un deuxième diagramme, appelé *diagramme entités-relations* (diagramme ER), beaucoup plus proche de SQL. Ce diagramme doit décrire la structure de la future base de données de façon aussi précise que possible. Chaque entité du MCD deviendra une entité du diagramme ER, puis une table de la base de données. Comme pour le MCD, on souligne les clés primaires et les entités sont reliées par des relations avec leurs cardinalités.

Il y a trois différences principales.

- 1) Les attributs du MCD sont nommés pour devenir des attributs dans une base de données et on précise leurs types. On peut également préciser les contraintes (valeur **NULL** interdite, par exemple).
- 2) On indique les clés étrangères, qui remplacent les relations présentées par un verbe à l'infinitif. Les cardinalités sont préservées.
- 3) En cas de relation *plusieurs à plusieurs*, i.e. de relation $(*, n)$ dans les deux sens, on crée une table *ad hoc* qui représente la relation. c'est le rôle de la table `relation_enseignements_cptces` utilisée pour modéliser le BUT.

Exemple. Voici le diagramme ER de l'exercice 2 :



Remarquez que dans le MCD, l'entité **Comptes** n'a pas d'attribut pour jouer le rôle de clé étrangère, car ce n'est pas une information inhérente à l'entité. On s'appuie sur le MCD et sur les cardinalités de la relation **Posséder** pour en déduire qu'une clé étrangère est nécessaire dans la table **comptes**. Autrement dit, la cardinalité $(1, 1) - (0, n)$ dans le MCD devient une clé étrangère du côté $(0, n)$.

Exercice 5. En vous aidant de l'exercice 1 et de l'exercice 3, établir le diagramme ER de la base de données qui modélise le BUT informatique. On précisera le type de chaque attribut, les cardinalités et les clés primaires et étrangères de chaque table.

Dans quel ordre faire les choses

Dans ce TP, on vous demande de construire le modèle conceptuel de données à *partir* d'une base de données existante. En pratique, on fait les choses dans le sens contraire : lorsque l'on doit concevoir une base de données, on commence par établir le modèle conceptuel de données et les diagrammes associés, et seulement ensuite on crée la base de données convenablement structurée.

Créer la table avant de réfléchir à sa conception est une *très mauvaise pratique* que vous devez bannir à tout jamais de vos habitudes.

2 Mise en pratique

Comme indiqué précédemment, en général on établit le MCD et le diagramme ER *avant* de créer les tables, contrairement à ce qui a été vu lors des TP précédents. Nous allons appliquer cette méthodologie sur un exemple.

2.1 Cahier des charges

À rendre

Dans cette partie, vous allez créer une base de données. Vous devrez déposer sur EPREL votre fichier `tp4_<nomDeFamille>.sql`, contenant l'ensemble des requêtes permettant de créer la table.

L'objectif est de concevoir une base de données pour la bibliothèque d'une petite ville. Les informations relatives aux adhérent-es sont : nom, prénom, date de naissance, adresse et le nombre maximal d'ouvrages empruntables (selon l'abonnement choisi). La ville est petite, donc il n'y a pas d'homonymes possibles : la donnée du nom et du prénom suffit à identifier chaque personne de manière unique. On enregistre également des informations sur les ouvrages disponibles : l'ISBN (identifiant international unique propre à chaque publication), titre, date de publication. Un même ouvrage peut être présent en plusieurs exemplaires.

Les informations sur les auteurs et autrices sont également stockées : nom, prénom, pseudonyme éventuel, date de naissance (si connue) et nationalité (si connue). Un auteur ou une autrice n'est présent-e en base que s'il ou elle a écrit au moins un ouvrage du catalogue. Un ouvrage peut avoir été écrit par 0 (si les auteurs et autrices ne sont pas identifiés), une ou plusieurs personnes. Enfin, on enregistre également les ouvrages en cours d'emprunt, avec pour chaque ouvrage l'identité de la personne qui l'a emprunté et la date de retour prévue. Si un ouvrage n'est pas en emprunt, ces informations ne sont pas renseignées.

2.2 À vous de jouer !

- 1) Quelles sont les entités et leurs attributs nécessaires pour modéliser la bibliothèque ?
- 2) En déduire diagramme du MCD, en précisant les relations et leurs cardinalité.
- 3) En se basant sur le MCD, construire un diagramme ER.
- 4) En se basant sur le diagramme ER, créer en SQLite la base de données.