

Conception des bases de données :

Modèle Entité-Association

La modélisation d'un problème, c'est-à-dire le passage du monde réel à sa représentation informatique, se définit en plusieurs étapes pour parvenir à son intégration dans un SGBD-R et permettre la manipulation des données par le langage SQL.

Classiquement, le processus de modélisation des données passe par deux phases :

- Réalisation d'un *modèle conceptuel*
- Traduction en un *modèle relationnel*

Le premier niveau de modélisation, dit conceptuel, consiste en une phase d'analyse du problème réel. Cette phase est assez délicate et permet de définir les données à utiliser, leur mode d'évolution dans le temps et les relations entre elles. C'est le moment où l'on se pose les questions essentielles comme celle de savoir à quel usage on destine le modèle informatique que l'on est entrain de constituer.

Ce travail est réalisé par des spécialistes de l'analyse. Il s'exprime dans un *formalisme* de type entité-association. Il existe d'autres types de formalismes comme le formalisme UML ou Merise (cf. ANNEXE).

Le second niveau de modélisation, dit relationnel, conduit à élaborer l'ensemble des objets manipulables par un SGBD-R. Ce travail est souvent réalisé par l'architecte de données, ou un administrateur de SGBD. Il peut être découpé en deux étapes :

- la conception de modèle *logique* (représentation en tables indépendantes du SGBD)
- la traduction en un modèle physique (propre à un SGBD spécifique). Tous les SGBD n'ont pas les mêmes caractéristiques du langage SQL.

1. Le Modèle Conceptuel de données (MCD)

Un schéma conceptuel représente :

- Les faits et les événements qui décrivent le monde à modéliser. Exemple : une compagnie aérienne, ses avions et ses pilotes.
- Certaines contraintes. Exemple : un pilote ne doit vole que s'il détient une licence en cours de validité et une qualification correspondant au type d'avion.

Les différentes techniques de MCD :

- Entité-Association (en anglais E-R pour *Entity Relationship*)
- Modèle binaire
- Modèle Z

Un MCD est un diagramme avec des entités et des associations que nous voyons dans la section suivante. Pour réaliser un tel digramme, il faut s'aider d'outils graphiques informatiques spécialisés (particulièrement dans le monde de l'entreprise). Ces mêmes outils génèrent le modèle logique (le plus souvent le modèle relationnel cf. séances 4 & 5 du cours) de la base. On se sert ensuite d'un script (le plus souvent SQL) pour créer les différentes tables et contraintes.

Quelques outils :

Avec licence	Libres
<ul style="list-style-type: none">- Microsoft Visio- Maestro SQL- Oracle Designer- PowerAMC (de Sybase)- ERwin- ER/Studio- ...	<ul style="list-style-type: none">- MySQL Workbench- DBDesigner

Le problème le plus couramment rencontré au sein des bases de données mal conçues est la redondance. Cette faute entraîne, à terme, des incohérences en modification, insertion et suppression de données, et rend la base peu performante.

2. Le modèle Entité-Association

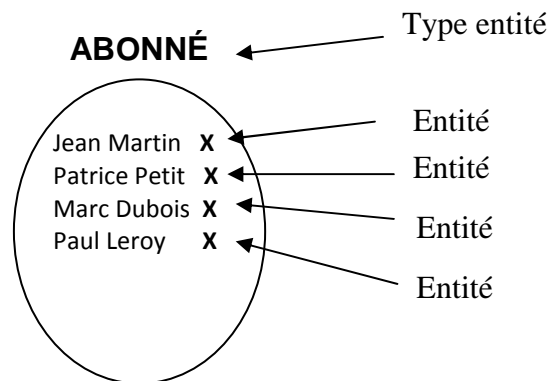
L'idée force du modèle entité-association est de représenter, par un schéma standardisé, les différents éléments constitutifs du système d'information, appelés *attributs* (exemple : nom, âge, ...), et les relations qui les unissent, appelées *associations*.

Une manière simple de modéliser est de décrire la réalité par une phrase :

Le sujet et le complément représentent des entités, et le verbe l'association.

Exemple : un utilisateur (entité) poste (association) une news (entité).

2.1. Entité



Ce schéma montre 4 entités du type-entité ABONNÉ.

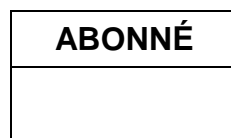
Définition : Une entité est un objet, une chose concrète ou abstraite qui peut être reconnue distinctement et qui est caractérisée par son unicité.

Il faut noter que les entités ne sont pas représentées sur un modèle entité-association.

Définition : Un **type-entité** désigne un ensemble de d'entités qui possèdent une sémantique et propriétés communes.

Les personnes, les livres, les voitures sont des type-entités. Dans le cas d'une personne, par exemple, les informations associées (ou propriétés) comme le nom, le prénom, ne changent pas de nature en fonction des personnes.

Dans un diagramme entité-association, un type-entité est représenté par un rectangle à cartouche. Le nom du type-entité est placé dans le cartouche : il est recommandé de choisir un nom commun décrivant le type-entité (exemple : ETUDIANT, ENSEIGNANT, MATIERE). Le nom du type-entité est écrit en majuscules dans le cartouche. Les attributs prendront place dans la partie basse.



Une entité est une occurrence ou instance de son type-entité. Par abus de langage, le terme entité est régulièrement utilisé pour désigner le type-entité et ses entités. Il ne faut cependant pas confondre ces deux notions.

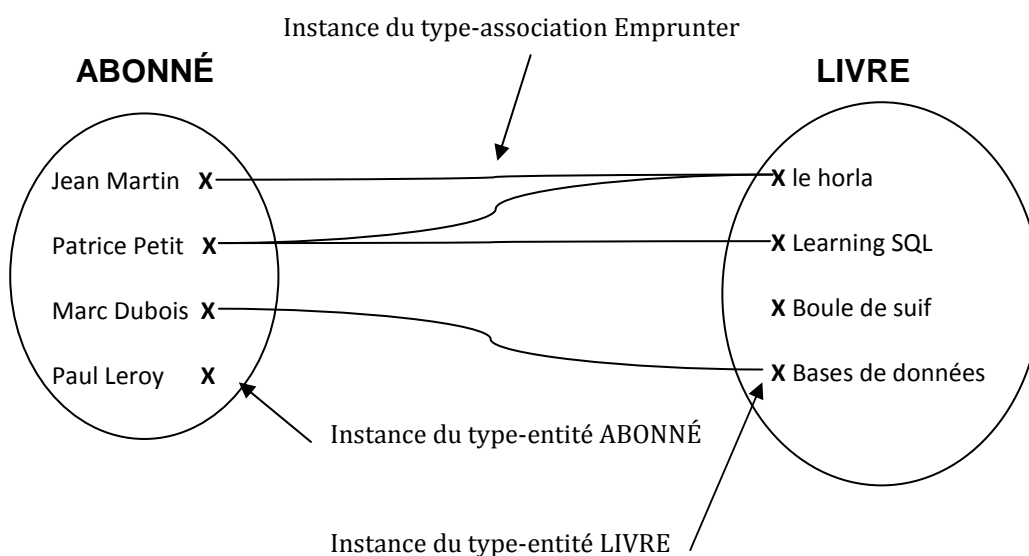
2.2. Association

Comme pour l'entité, le terme d'association est utilisé pour désigner un type-association.

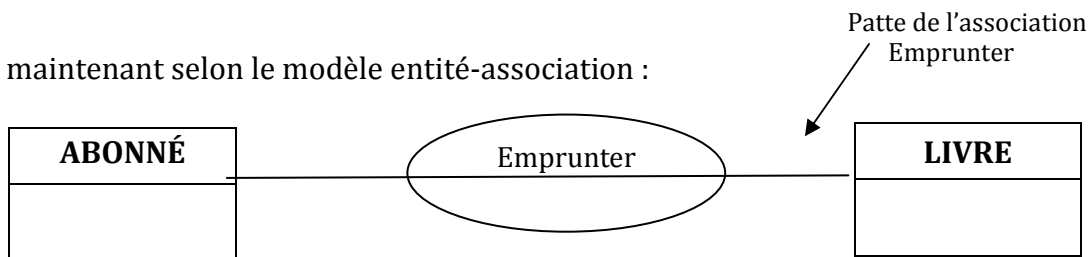
Définition : Une **association** (ou relation) est un lien entre plusieurs entités.

Exemple : L'ABONNÉ Paul Leroy Emprunte le LIVRE « Le horla ».

Nous représentons ce lien de manière informelle, ci-dessous :



Et maintenant selon le modèle entité-association :



Les propriétés ne sont pas représentées.

Dans le modèle entité-association, le type-association est représenté par une ellipse à cartouche. Le nom du type-association est placé dans le cartouche. Ce nom doit être un verbe à l'infinitif, à la forme passive ou bien accompagné d'un adverbe. Exemple : Enseigner, AvoirLieuDans. Notez que l'initiale du nom de l'association est en majuscule.

Les attributs de l'association sont placés dans la partie inférieure au cartouche.

Définition : Les type-entités intervenant dans un type-association sont appelés **participants**.

Définition : L'ensemble de participants d'un type-association est appelé la **collection** de type-association. La collection d'un type-association peut ne comporter d'un seul type-entité.

Définition : La dimension ou l'arité d'un type-association est le nombre de type-entités contenu dans la collection.

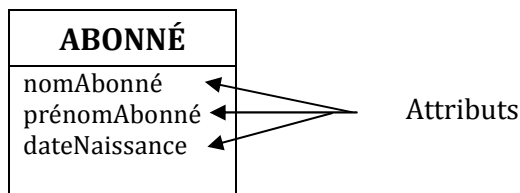
Un type-association est dit :

- N-aire : dans le cas général,
- Binaire : dans le cas où $n=2$ (la collection contient deux type-entités),
- Ternaire : dans le cas où $n=3$.

Les type-associations avec $n>2$, bien qu'existants, sont rares et problématiques.

2.3. Attribut et valeur

Définition : Un **attribut** (ou propriété) est une caractéristique associée à un type-entité ou à un type association.



Le nom des attributs est composé de deux morphèmes accolés : le premier tout en minuscule et la première lettre du second en majuscule.

Définition : Une **valeur** au niveau du type-entité ou de type-association, chaque attribut possède un domaine qui définit l'ensemble des valeurs possibles qui peuvent être choisies pour lui (entier, chaîne de caractères, booléen...). Au niveau de l'entité, chaque attribut possède une **valeur** compatible avec son domaine.

Règles :

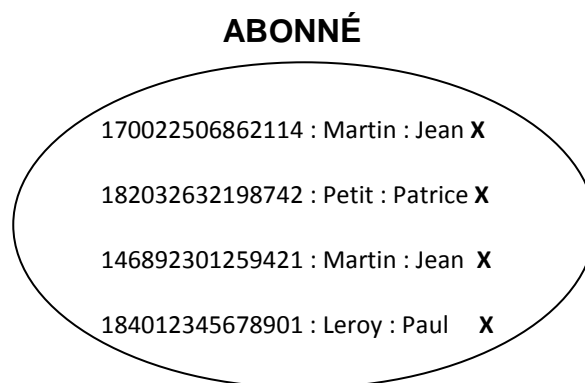
- Dans le modèle entité-association, chaque attribut est destiné à recevoir une valeur.
- Un attribut ne peut être partagé par plusieurs types-entité ou type-associations.
- Un attribut est une donnée élémentaire, ce qui exclut les données calculées ou dérivées.
- Un attribut peut être placé dans un type-association uniquement lorsqu'il dépend de toutes les entités liées par le type-association.

Un type-association peut ne pas posséder d'attribut explicite et cela est relativement fréquent, mais nous verrons qu'il possède au moins des attributs implicites.

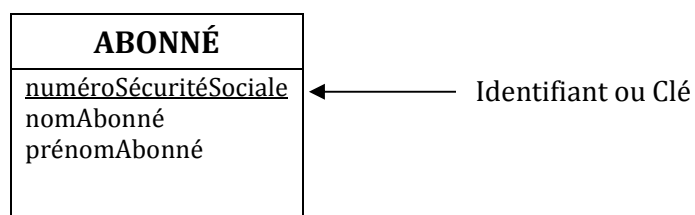
2.4. Identifiant

Définition : Un **identifiant** (ou **clé**) d'un type-entité ou d'un type-association, est un ensemble minimal d'attributs qui permet d'identifier chaque entité ou association de manière unique.

Par exemple dans le cas du type-entité ABONNÉ, si nous choisissons nomAbonné et prénomAbonné comme identifiants(ou clés) et que l'on a des homonymes dans la liste des abonnés, il y aura doublons dans les clés choisies ce qui est interdit. Par contre si nous prenons comme clé une entité numéroSecuritéSociale, nous aurons la garantie de ne pas avoir de doublons. Deux abonnés peuvent avoir le même nom et prénom mais pas le même numéro de sécurité sociale.



Un identifiant est représenté comme suit dans le modèle entité-association :



Nous pouvons donc voir qu'un identifiant est représenté dans ce modèle comme un attribut à la seule différence qu'il est souligné. Ce qui permet de le distinguer des autres attributs. Les identifiants sont également placés en tête des autres attributs.

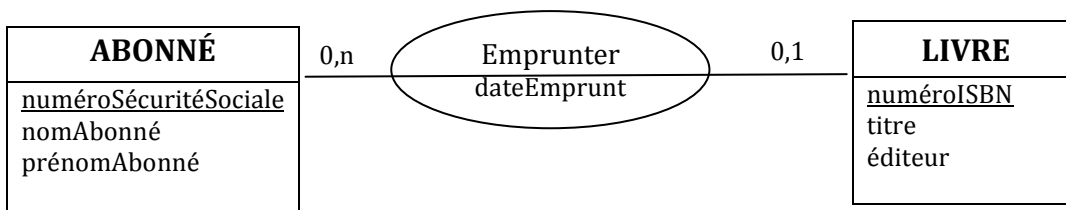
Règle :

- Chaque type-entité possède au moins un identifiant, éventuellement formé de plusieurs attributs.
- Tout type-entité, et type-association possède au moins un identifiant.
- Il ne peut y avoir la même valeur d'identifiants pour deux instances d'un type-entité (ou type-association).

2.5. Cardinalités d'un type-association

Définition : Les **cardinalités** d'une *patte* reliant un type-association et un type-entité précisent le nombre de fois minimal et maximal d'interventions d'une entité du type-entité dans une association de type-association. La cardinalité minimale doit être inférieure ou égale à la maximale.

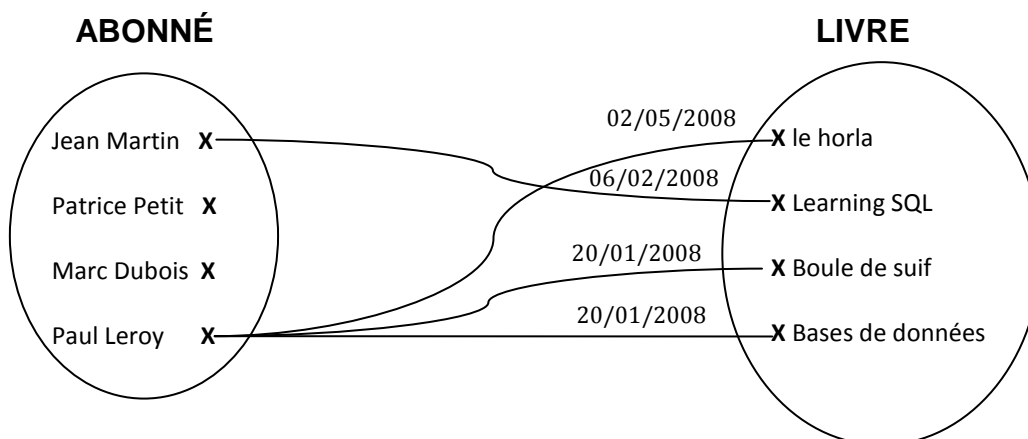
Exemple :



Il faut lire la cardinalité comme suit :

- Un abonné peut Emprunter plusieurs livres.
- Un livre ne peut être Emprunté que par un seul abonné.

Traduction en modèle informel (ou une instance possible du modèle ci-dessus):



Règles :

- L'expression de la cardinalité est obligatoire pour chaque patte d'un type-association.
- Il ne peut y avoir de cardinalité maximale égale à 0, car elle rendrait le type-association inutile.
- Une cardinalité minimale est toujours 0 ou 1 et une cardinalité maximale est toujours 1 ou n.
- Si une cardinalité est connue et vaut 2 ou plus, alors nous considérons qu'elle est indéterminée et vaut n. En effet, si cette valeur est définie lors de la conception, il se peut qu'elle évolue dans le futur. Il est donc considéré n comme inconnue dès la conception.
- Les seules cardinalités admises sont :
 - 0,1 : une occurrence du type-entité peut exister en étant impliquée soit dans aucune association soit au maximum dans une seule.
 - 0,n : une occurrence du type-entité peut exister en étant impliquée soit dans aucune association soit dans plusieurs associations (sans limite).
 - 1,1 : une occurrence du type-entité ne peut exister que si elle est impliquée dans exactement une association.
 - 1,n : une occurrence de type-entité ne peut exister que si elle est impliquée dans au moins une association.

2.6. Identifiant d'un type-association

Les type-associations peuvent être catégorisés en fonction des cardinalités maximales de leurs pattes :

Catégories de Type-association	Description	Autre appellation
Binaire	La cardinalité maximale sur chacune de ses pattes est 1.	Type-association 1 vers 1
Fonctionnel	Une cardinalité maximale est à 1 et une cardinalité maximale est à n.	Type-association 1 vers n
maillé	La cardinalité maximale de chacune de ses pattes est à n.	Type-association n vers n

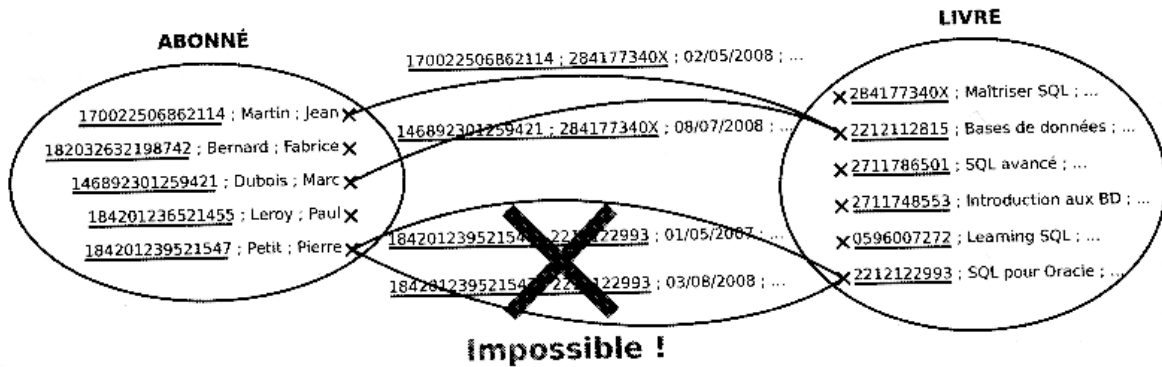
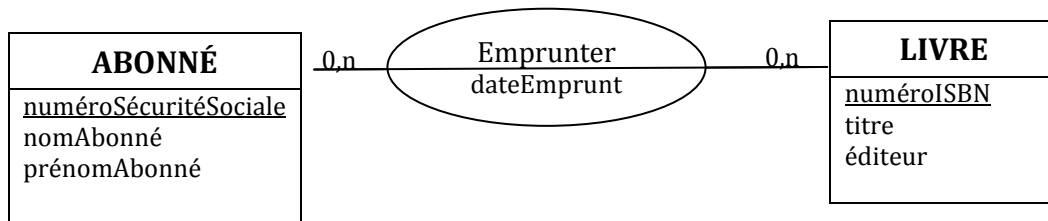
Seuls les type-associations maillés ont des propriétés.

Règles :

- L'identifiant d'un type-association maillé est composé des identifiants de ses participants. Cet identifiant n'est pas mentionné sur le modèle, il est *implicite*.

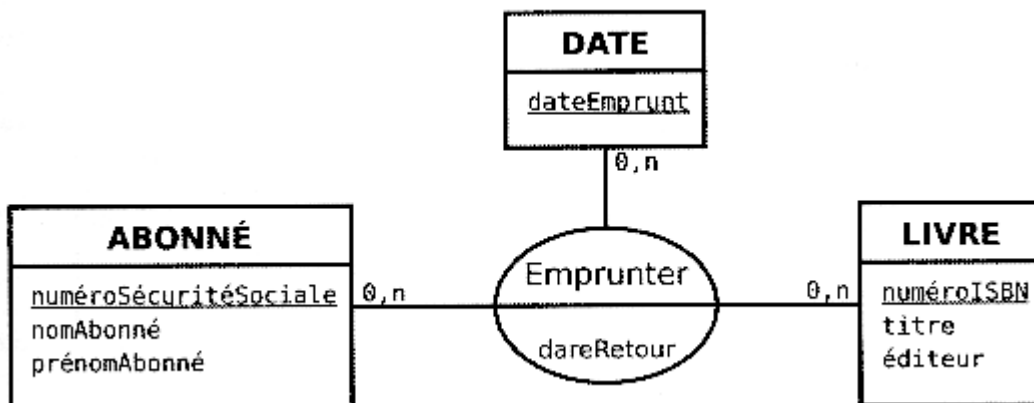
Explications :

Soit le modèle suivant. Il précise qu'un livre peut être Emprunté par plusieurs abonnés, et qu'un abonné peut Emprunter plusieurs livres.

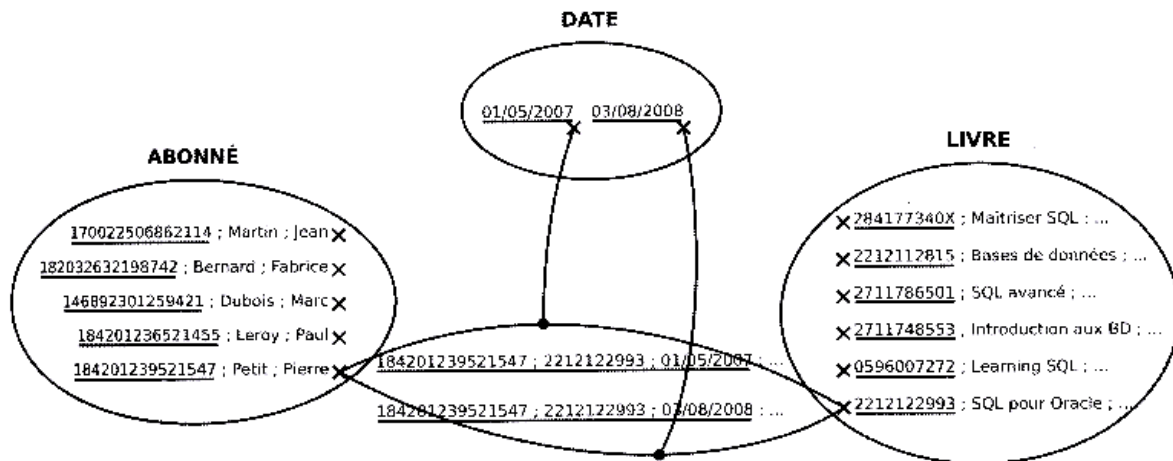


Si un abonné Emprunte un même livre à deux différentes dates, alors l'identifiant du type-association Emprunter comptera un doublon, ce qui est impossible.

Pour permettre à notre modèle d'autoriser un tel emprunt, il faut rajouter un 3^{ème} identifiant dateEmprunt au type-association Emprunter. Pour cela il faut rajouter un type-entité DATE avec comme identifiant dateEmprunt. Ainsi nous obtenons le modèle entité-association suivant :

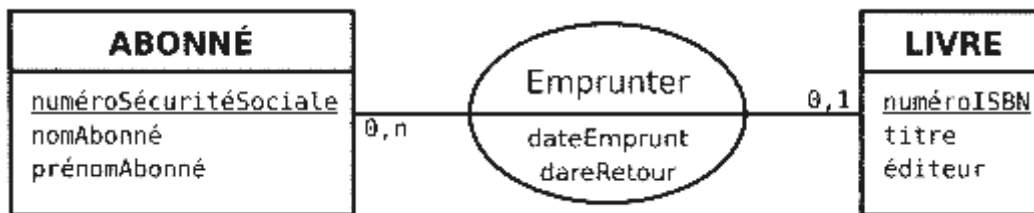


Ce modèle autorise donc l'emprunt d'un même livre par un abonné à des différentes dates.

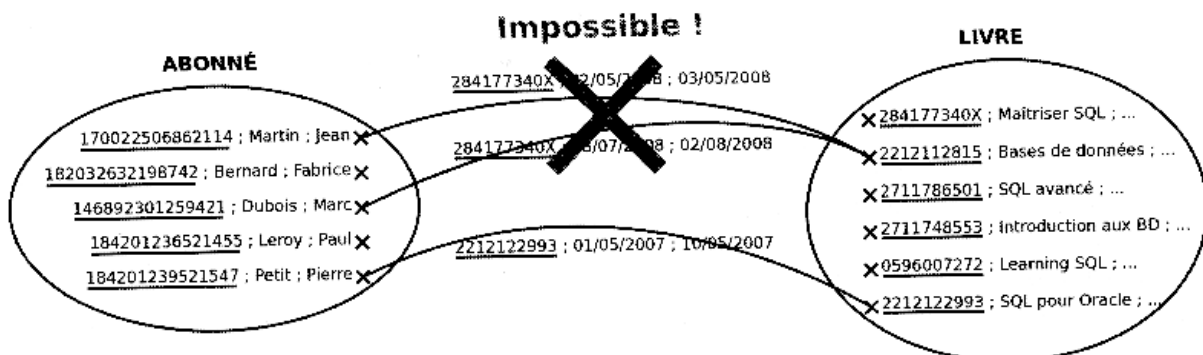


- L'identifiant d'un type-association fonctionnel est composé des identifiants du participant situé du côté de la cardinalité maximale de 1. Cet identifiant n'est pas mentionné sur le modèle, il est *implicite*.

Exemple : dans le modèle ci-dessous, un abonné peut emprunter autant de livres qu'il le souhaite. Par contre un livre ne peut être emprunté que par un abonné.



D'après la règle d'un type-association fonctionnel, l'identifiant du type-association est celui du type-entité de cardinalité maximale est égale à 1. Soit donc l'identifiant numéroISBN. Or si Emprunter a pour clé (ou identifiant) numéroISBN, lorsqu'un abonné emprunte un livre une deuxième fois, nous aurons deux fois la même occurrence de numéroISBN. Ce qui est impossible. Le modèle donné ici est donc inadapté à la réalité.



- Un type-association binaire dit 1 vers 1, possède deux identifiants candidats issus des deux participants. Il suffit d'en choisir un. Cet identifiant n'est pas mentionné sur le modèle, il est *implicite*.

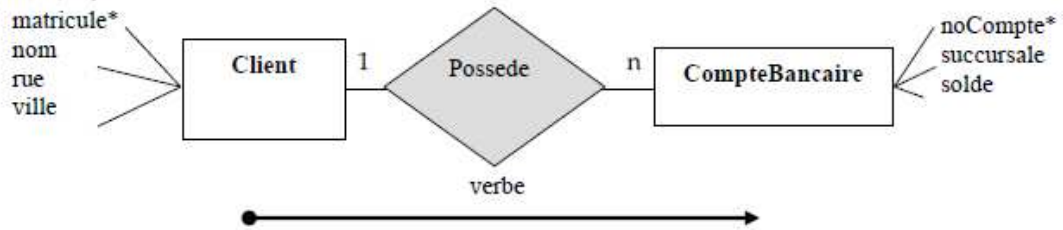
Dans un tel cas, l'abonné ne peut emprunter qu'un livre maximum, et un livre ne peut être emprunté que par un seul abonné. Ce qui est assez limité comme mode d'emprunt. Une telle cardinalité du modèle n'est pas adapté à la réalité.

ANNEXE

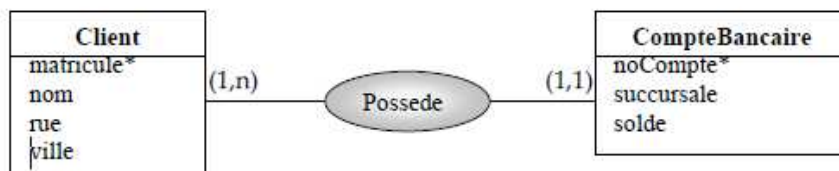
Les différents formalismes de la modélisation de données :

Le formalisme de représentation des modèles a évolué pour donner naissance à une panoplie de langages graphiques de modélisation.

(Chen : E/A)



(Merise)



(UML)

