

Chapitre 2

Modèle Conceptuel de données

I. Introduction

Le **modèle conceptuel de données** MCD (ou **modèle entité-association** MEA, ou **Entity-Relationship Model** en anglais) a été introduit dans les années 70 comme une amélioration du modèle relationnel introduit par Codd en 1970: le MCD est plus facile à lire pour la construction de bases de données. Il devient le modèle le plus utilisé pour représenter dans un premier temps la structure de données. Actuellement, il n'y a pas de standard Entité-Association, il existe une large variété de notations et de concepts.

Deux utilisations des schémas EA :

- Concevoir le schéma de la base de données avant que la moindre donnée ne soit stockée : c'est la conception conceptuelle. A partir d'un tel schéma, il est plus clair de raisonner sur les concepts à stocker, les attributs dont on a besoin, les relations entre les concepts... Une fois terminée, un tel schéma est traduit selon des règles précises en son schéma relationnel correspondant. La construction du schéma entité-association ainsi que son passage au modèle relationnel seront vus dans les chapitres 3 et 4.
- Faire de la réingénierie : à partir d'une base existante complexe, dont la documentation est inexistante, on peut créer un schéma EA pour avoir une première vue synthétique de la situation et prendre des décisions comme restructurer la base, la normaliser, en créer une nouvelle plus large permettant de stocker les données de l'ancienne,...

II. Concepts de base

1. **Modèle entité-association (EA)**

Le modèle EA propose une description sur la base des trois concepts de base qui sont l'identification des objets, des liens entre ces objets et des propriétés de ces objets :

- objet → entité
- lien → association
- propriété → attribut

2. **Définitions**

a. **Entités**

Une **entité** est un objet concret ou abstrait du monde réel à propos duquel on veut enregistrer des informations.

ex : *M. Dupont, Mme Dupont, un crayon, l'atelier de distribution, le bureau du directeur...*

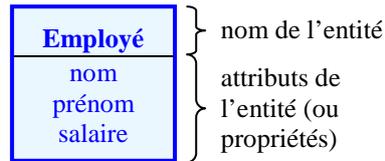
Un **type d'entité (TE)** est un ensemble d'entités qui possèdent les mêmes caractéristiques.

ex : *Personne (représentation de l'ensemble des personnes telles que les entités M. Dupont et Mme Dupont), Employé (représentation de l'ensemble des employés), Article (représentation de l'ensemble des articles tels que les crayons et les livres)...*

Un **attribut d'une entité** est une propriété associée à un TE. L'ensemble des attributs d'un TE représente l'ensemble des informations inhérentes que l'on souhaite conserver sur les entités du TE.

ex : *(nom, prénom, salaire) sont des attributs du TE Employé...*

Une entité munie de ses attributs sera modélisée de la manière suivante :



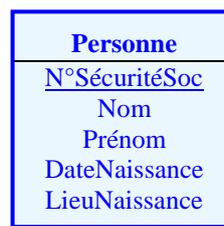
On distingue deux types d'entités :

- **Entité forte** = entité qui n'a pas besoin d'une autre entité pour exister.
- **Entité faible** = entité qui a besoin d'une autre entité pour pouvoir être définie. Elles sont notées avec un double rectangle. On distingue deux cas :

o 1^{er} cas :

Employé est une entité faible car l'ensemble des *numéros de sécurité sociale* des élèves est contenu dans l'ensemble des *numéros de sécurité sociale* des personnes.

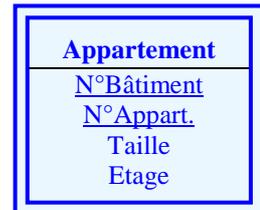
On dit qu'il y a un **lien d'héritage** entre les deux entités.



o 2nd cas :

Appartement est une entité faible car sa clé (on définira cette notion par la suite) est composée de l'attribut clé de *Bâtiment* (*N°Bâtiment*) et d'un autre attribut (*N°Appart.*).

On dit qu'il y a un **lien d'identification** entre les deux entités.



b. Associations

Une **association** (ou une **relation**) est un lien entre plusieurs entités, où chaque entité liée joue un rôle déterminé ; si l'association lie des entités du même type, elle est dite **cyclique** ou **réflexive** et, dans ce cas, la spécification des rôles est indispensable.

ex : *lien de travail entre l'employée M. Durant et le service de gestion, lien de mariage entre M. X et Mme Y.*

Un **type d'association (TA)** est un ensemble d'associations qui possèdent les mêmes caractéristiques (liant des entités de mêmes types avec mêmes rôles et mêmes propriétés, respectifs)

ex : *le TA « fabrique » lie le TE « Atelier de fabrication » au TE « Article », le TA « travaille » lie « Employé » à « Service », le TA « est-marié-avec » lie « Personne » à lui-même.*

Un **attribut d'une association** est une propriété associée à un TA. L'ensemble des attributs d'un TA représente l'ensemble des informations inhérentes que l'on souhaite conserver sur les associations du TA.

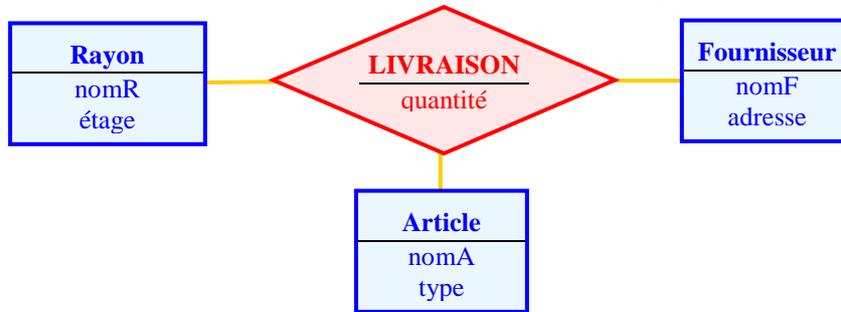
ex : *(quantité-en-fabrication) est un attribut du TA fabrique...*

Une association sera modélisée de la manière suivante :

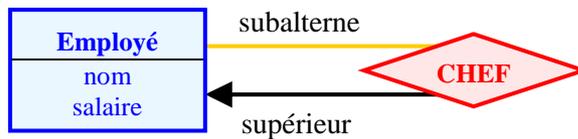
- **Association binaire** : C'est une association qui relie deux entités.



- **Association n-aire** (ici, avec $n = 3$) : C'est une association qui relie n entités différentes.



- **Association réflexive** : C'est une association qui agit sur la même entité, mais avec des rôles différents.



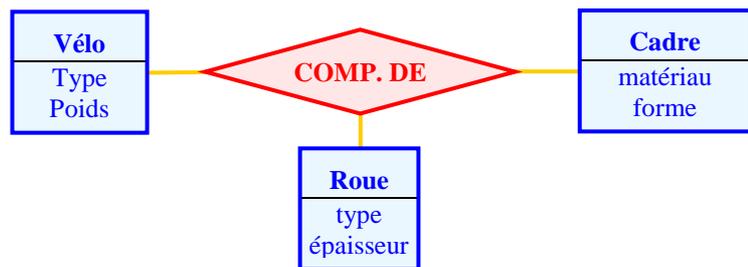
Le nom des associations peut-être (entre autres) :

- Un verbe (plutôt pour une association binaire ou réflexive) pour décrire des rôles réciproques entre deux entités ;
- Un nom (plutôt pour les associations n -aires) pour représenter une entité complexe définie à partir d'autres entités.

Les associations peuvent avoir des significations différentes :

- **Composition** :

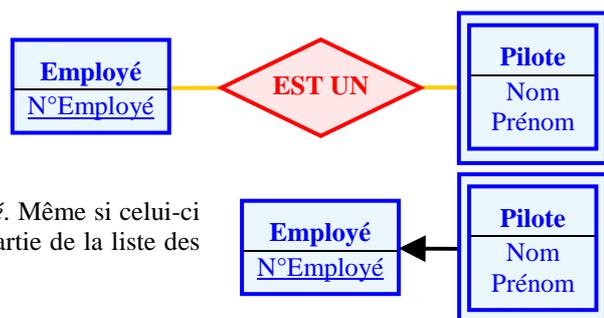
Le sens de l'association est de définir l'entité *Vélo* comme composé d'autres entités (*Roue* et *Cadre*). Ce type d'association est à utiliser avec précaution car on peut modéliser, dire tout et n'importe quoi avec ce lien de composition.



- **Héritage** :

Pour représenter ce lien d'héritage, on peut utiliser les deux notations ci-contre.

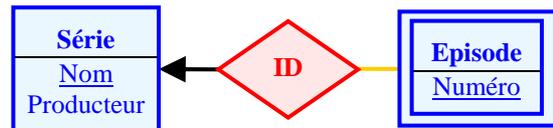
Quelle que soit la notation, la clé de *Pilote* est son attribut *N°Employé*. Même si celui-ci n'est pas noté, il fait quand même partie de la liste des attributs de *Pilote*.



– **Identification :**

Pour représenter ce lien d'identification, on utilise la notation ci-contre.

La clé de *Episode* est son couple d'attributs (*Nom*, *Numéro*). Même si l'attribut *Nom* n'est pas noté, il fait quand même partie de la liste des attributs de *Episode*.



– **Muni d'attributs :**

On peut ajouter des attributs propres à une association, ils dépendent de toutes les entités reliées par cette association.

Prix est un attribut propre à l'association *Catalogue* : le *Prix* dépend à la fois du *Fournisseur* et du *Produit*.

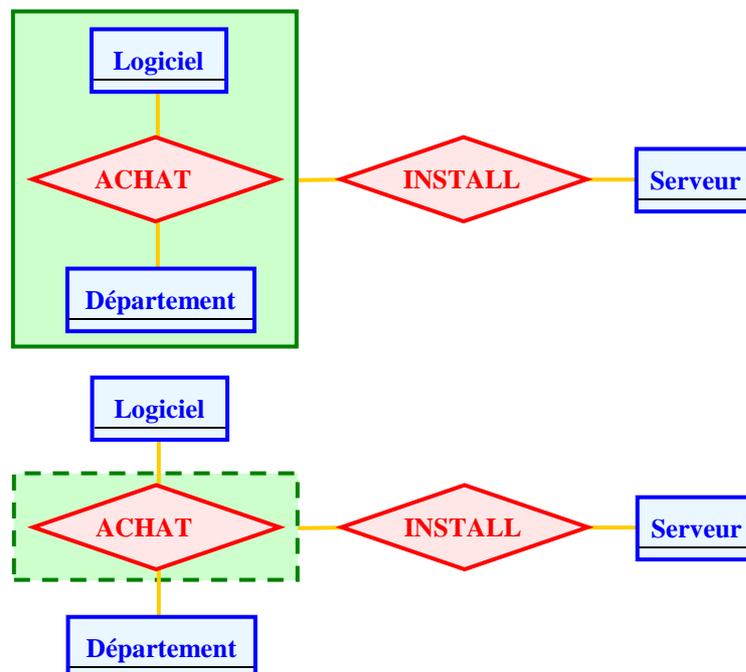


– **Associations considérées comme entités :**

Certaines associations peuvent être vues comme des entités, pour notamment pouvoir construire des associations d'associations.

Exemple :

Supposons que l'on ait à représenter un ensemble de *départements universitaires* avec leurs *serveurs* et les *logiciels* qui y sont installés. On veut représenter le fait qu'un *département* qui a acheté un *logiciel* l'a forcément installé sur un *serveur*. Avec une association ternaire (3-aire), ou pourrait représenter les choses ainsi de la manière. Malheureusement, les cardinalités ne nous permettent pas de spécifier que pour un couple (*logiciel*, *département*), il existe obligatoirement au moins 1 *serveur*. Pour cela, il convient de créer une association supplémentaire que l'on utilisera comme entité.



c. Occurrence et population

On appelle **occurrence d'un TE (TA)** toute entité (association) appartenant à l'ensemble décrit par le TE (TA). Une occurrence de TE est un ensemble de valeurs (éventuellement vide) pour chaque attribut du TE. Une occurrence de TA est un ensemble de valeurs (éventuellement vide) pour chaque attribut du TA et chaque occurrence de Tes qui joue un rôle dans le TA.

On appelle **population d'un TE (TA)** l'ensemble des occurrences du TE (TA).

III. Diagrammes Entité-Association

1. Diagramme

Le modèle entité-association permet une représentation graphique assez lisible du schéma d'une BD, appelée diagramme entité-association (diagramme EA)

Les TE sont représentés par des rectangles, les TA par des hexagones ou autres symboles similaires (ovale, losange, ...), les attributs sont soit rattachés au TE (TA) par des traits, et représentés par des ovales, soit listés à l'intérieur du rectangle (hexagone), au dessous du nom du TE (TA) et séparés de celui-ci par une barre.

2. Contraintes de cardinalité

Les **cardinalités** d'un lien entre un TE et un TA sont le **minimum** (min) et le **maximum** (max) de fois qu'une entité peut intervenir dans une association de ce type.

Formellement : si un TA A lie n TEs (E_1, E_2, \dots, E_n) et si (a_i, b_i) représente les cardinalités du lien entre E_i et A , alors a_i et b_i décrivent respectivement le nombre minimum et le nombre maximum de $(n-1)$ -uplets de $E_1 \times \dots \times E_{i-1} \times E_{i+1} \times \dots \times E_n$ auxquels une instance quelconque de E_i est reliée.

Remarque :

- La cardinalité maximum b_i est toujours ≥ 1 .
- Si une cardinalité est connue et vaut $k > 1$, on considère parfois qu'elle est indéterminée N , de telle sorte qu'on ne puisse avoir que des cardinalités 0, 1 ou N .
- Il est impossible d'avoir un couple de cardinalités (a, b) avec $a > b$.

Propriétés :

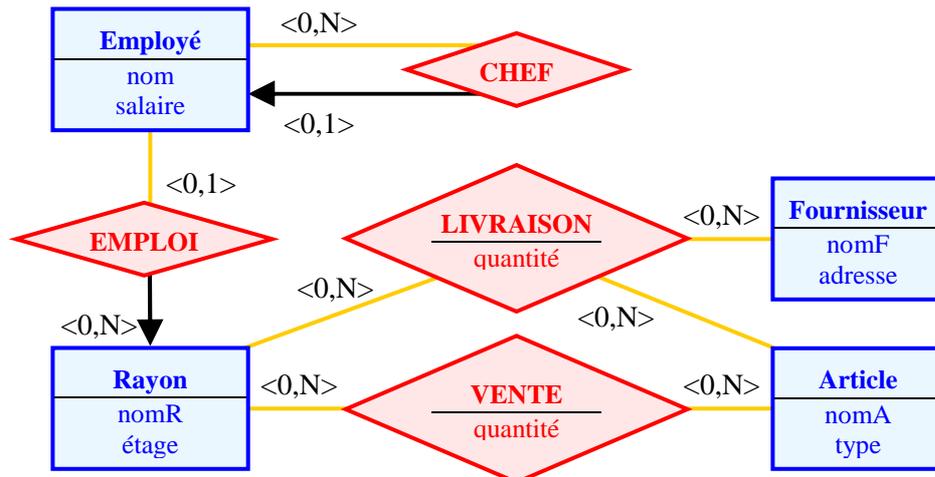
- $min = 0$: toute entité peut exister tout en étant impliquée dans aucune association ; on parle de **participation partielle** du TE au TA.
- $min = 1$: aucune entité ne peut exister sans être impliquée dans au moins une association ; on parle de **participation totale** du TE au TA.
- $min = N$: aucune entité ne peut exister sans être impliquée dans plusieurs associations (cas rare et complexe).
- $max = 0$: toute entité ne peut pas être impliquée dans une association (cas à éviter)
- $max = 1$: toute entité peut être impliquée dans au plus une association.
- $max = N$: toute entité peut être impliquée dans plusieurs associations.

Classification

Classification d'une association binaire A selon les cardinalités maximums des entités participantes E et F :

- A est **1:1** si $maxCard(E, A) = 1$ et $maxCard(F, A) = 1$
- A est **1:N** si $maxCard(E, A) = 1$ et $maxCard(F, A) = N$ (ou vice-versa)
- A est **M:N** si $maxCard(E, A) = N$ et $maxCard(F, A) = N$

Exemple :



Dans l'exemple ci-dessus, expliquons les contraintes de cardinalité qui entourent l'association :

- **EMPLOI :**

EMPLOI		
Employé	Rayon	Commentaires
e ₁	r ₁	Un employé e ₁ peut être employé dans aucun rayon (minimum = 0) et peut travailler dans un seul rayon r ₁ au maximum (maximum = 1), ce qui explique la contrainte de cardinalité <0,1> entre <i>Employé</i> et <i>EMPLOI</i> .
e _±	r _±	
e ₁	r ₁	Dans un rayon, on peut avoir aucun employé qui y travaille (minimum = 0) mais on peut en avoir autant que l'on veut (maximum = N), ce qui explique la contrainte de cardinalité <0,N> entre <i>Rayon</i> et <i>EMPLOI</i> .
e ₂	r ₁	
...	r ₁	

- **LIVRAISON :**

LIVRAISON			
Fournisseur	Rayon	Article	Commentaires
f ₁	r ₁	a ₁	Un fournisseur f ₁ peut être faire l'objet de la livraison d'aucun article dans aucun rayon (minimum = 0). De plus, un fournisseur f ₁ peut fournir plusieurs articles (a ₁ , a ₂ ,...) dans plusieurs rayons (r ₁ , r ₂ ,...) différents (maximum = N), ce qui explique la contrainte de cardinalité <0,N> entre <i>Fournisseur</i> et <i>LIVRAISON</i> .
f ₁	r ₁	a ₂	
f ₁	r ₂	a ₂	
f ₁	r ₂	a ₁	
f ₁	
f ₁	r ₁	a ₁	Un rayon r ₁ peut être faire l'objet de la livraison par aucun fournisseur ou d'aucun article (minimum = 0). De plus, un rayon r ₁ peut être rempli grâce à plusieurs articles (a ₁ , a ₂ ,...) et par plusieurs fournisseurs (f ₁ , f ₂ ,...) différents (maximum = N), ce qui explique la contrainte de cardinalité <0,N> entre <i>rayon</i> et <i>LIVRAISON</i> .
f ₁	r ₁	a ₂	
f ₂	r ₁	a ₂	
f ₂	r ₁	a ₁	
...	r ₁	...	
f ₁	r ₁	a ₁	Un article a ₁ peut être faire l'objet de la livraison par aucun fournisseur ou dans aucun rayon (minimum = 0). De plus, un article a ₁ peut être fourni par un seul fournisseur f ₁ mais dans plusieurs rayons (r ₁ , r ₂ ,...) différents (maximum = N), ce qui explique la contrainte de cardinalité <0,N> entre <i>article</i> et <i>LIVRAISON</i> .
f ₁	r ₂	a ₁	
f _±	r _±	a _±	
...	...	a ₁	

3. Cardinalités d'un attribut

Les **cardinalités d'un attribut** spécifient le nombre de valeurs de cet attribut qui sont autorisées (au minimum, au maximum) dans :

- une occurrence du TE (TA), si l'attribut est directement rattaché à un TE (TA).
- une valeur de l'attribut dont il est composant, sinon.

Un attribut est dit :

- **simple** s'il n'est pas décomposé en d'autres attributs : ses valeurs sont atomiques; un domaine lui est associé ;
- **complexe** s'il est décomposé en d'autres attributs; ses valeurs sont composées ;

- **monovalué** s'il ne peut prendre qu'une seule valeur par occurrence ($maxCard = 1$) ;
- **multivalué** s'il peut prendre plusieurs valeurs par occurrence ($maxCard > 1$) ;
- **obligatoire** s'il doit prendre une valeur au moins par occurrence ($minCard = 1$) ;
- **facultatif** s'il peut ne pas prendre de valeur dans une occurrence ($minCard = 0$).

4. Identifiants des TE et TA

Un **identifiant** d'un TE (TA) est constitué par un ou plusieurs attributs qui doivent avoir une valeur unique pour chaque entité (association) de ce type. Les attributs appartenant à un identifiant doivent être de cardinalité (1,1). Un TE (TA) peut avoir aucun ou plusieurs identifiants :

ex : n°employé et (nom+prénoms) sont deux identifiants du TE Employé, si dans cette entreprise il n'y a jamais deux employés ayant les mêmes nom et prénoms, ou le même numéro

Les identifiants (ou clés primaires) peuvent être représentés sur le diagramme en les soulignant. Pour éviter des surcharges, les identifiants des TA peuvent être précisés textuellement en commentaire du diagramme.

Règles :

- si le TA a des cardinalités (1,1) pour un des TE liés, alors tout identifiant de ce TE est identifiant du TA.
- si plusieurs occurrences du TA lient les mêmes occurrences des TE, alors chaque identifiant du TA contient au moins un attribut du TA.

TE faible :

Un TE E est dit **faible** si aucun sous-ensemble de ses attributs ne constitue un identifiant et qu'un identifiant peut être défini en intégrant un identifiant d'un autre TE qui lui est lié par un TA binaire (**d'identification**) de cardinalité (1,1) pour E .

L'identifiant d'un TE faible (qui est le même que celui du TA d'id.) est constitué de l'identifiant du TE dont il dépend et d'un (ou plusieurs) attributs du TE faible

IV. Contraintes d'intégrité

1. Définition

Une **contrainte d'intégrité** (CI) est un ensemble de règles de contrôle de cohérence des valeurs prises par les différentes entités et associations d'une BD et qui ne peut pas être décrite avec les concepts du modèle.

ex : Si une personne participe à l'association Mariage, alors la valeur de son attribut « état civil » doit être « mariée » :

$$\forall x,y \in \text{Personne}, \langle x,y \rangle \in \text{Mariage} \Rightarrow x.\text{état-civil} = \text{'marié'}$$

ex : Seuls les hommes peuvent participer à l'association mariage dans le rôle « époux »

$$\forall x,y \in \text{Personne}, \langle \text{époux}:x,y \rangle \in \text{Mariage} \Rightarrow x.\text{sexe} = \text{'M'}$$

2. Contraintes d'intégrité sur les attributs

Les CI les plus fréquentes limitent les valeurs d'un attribut à certaines valeurs du domaine sous-jacent.

ex : Age $\in [0, 130]$.

Les limites peuvent être définies en fonction du contexte (valeur d'un autre attribut, participation à une association, ...)

ex : Si mois $\in \{4,6,9,11\}$ alors jour $\in [1:30]$, sinon si mois = 2 alors jour $\in [1:29]$ sinon jours $\in [1:31]$.

ex : Si une personne participe à l'association Mariage, alors la valeur de son état civil doit être « marié ».