

[BDDr] Du modèle Entité-Association (partie 2) au Modèle Relationnel

ESIEE Paris

Bases de Données Relationnelles

Lionel VILLARD

Notation simplifiée pour les Type-Associations de spécialisation

Notions avancées sur le modèle Entité-Association

- Bien choisir l'identifiant

- Des noms uniques

- Quand fusionner ou supprimer ?

- Normalisation des Type-Entités et des Type-Associations

 - Rappel : la clé

 - Première forme normale (1FN)

 - Deuxième forme normale (2FN)

 - Troisième forme normale (3FN)

 - Forme normale de Boyce-Codd (BCNF)

Synthèse : Etapes de conception d'un modèle Entité-Association

Introduction au modèle relationnel

- Introduction

- Définitions du Modèle Relationnel

 - Relation, Attribut et Domaine

 - Tuple

 - Degré et cardinalité

 - Les clés (candidate, primaire, étrangère)

 - Synthèse

 - Base de données et schéma de relation

 - Le schéma relationnel d'une base de données

- Passer du modèle Entité Association au Modèle Relationnel

 - Règles Générales

 - Cas particuliers des Type-Associations 1:1

 - Vidéo

TD 1 / Conception et modèle Entité-Association

- Bibliothèque partie 1 :
 - [https://esiee.blackboard.com/webapps/blackboard/content/listContent.jsp?course_id= 4954 1&content id= 37422 1&mode=reset](https://esiee.blackboard.com/webapps/blackboard/content/listContent.jsp?course_id=4954_1&content_id=37422_1&mode=reset)
 - Notation simplifiée pour les **Type-Associations de spécialisation (héritage)**
 - Un **type entité Date**

Un même abonné doit pouvoir emprunter un même exemplaire à des dates différentes. Il est donc possible de faire modéliser un Type-Entité concernant le temps pour ce Type-Association. Ainsi, le Type-Entité **DATE** permettra pour chaque exemplaire emprunté par l'abonné, de savoir s'il l'a été emprunté plusieurs fois (et à quelle date il a été emprunté).

Pourquoi des règles de bonnes pratiques sur le modèle Entité-Association ?

- Eviter les sources d'**incohérences** et de **redondances** qui pourront, ensuite, se manifester au niveau de la base de données.
- Les règles de conception (normalisation) sont des méthodes complémentaires « **une seule place pour chaque fait** »
- Il peut exister, occasionnellement et en particulier en fonction de besoins spécifiques, de **bonnes raisons de ne pas appliquer ces règles**.

Choix des identifiants

Il faut éviter de choisir un identifiant qui :

- soit susceptible de **changer en fonction du temps** (*nécessitant de changer d'identifiant car n'étant plus pertinent ou impossible, e.g. un id construit sur un numéro de salle, alors qu'avec la construction d'un nouveau bâtiment des nouvelles salles pourraient avoir le même numéro*).
- soit composé de plusieurs attributs pour identifier un Type-Entité car :
 - ils dégradent les performances du SGBD (*mais index*)
 - l'unicité finit généralement par être démentie (*e.g. une année pour identifier un recensement quadriennal de l'INSEE mais qui pourrait devenir un jour semestriel*).

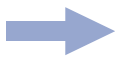
Choix des identifiants

Personne
<u>Nom</u>
<u>Prénom</u>
DateNaissance

Ville
<u>codePostal</u>
NomVille
NombreHabitants

Exemples d'identifiants qui semblent s'imposer mais qui posent problème :

- Le **couple (nom, prénom)** : les homonymes sont nombreux;
- Le **code postal** ne constitue pas un bon identifiant : une commune (la ville) peut avoir plusieurs codes postaux (et il y a des codes postaux qui peuvent recouvrir plusieurs communes).

- 
- Il faut donc définir un **identifiant arbitraire** (un entier) pour les Types-Entités.
 - Il deviendra une **clé primaire** dans le modèle relationnel et auto-incrémentée par le SGBD.

Des noms uniques

Dans le modèle Entité-Association, les noms d'un Type-Entité, d'un Type-Association ou d'un attribut, doivent être **uniques**.

Si ce n'est pas le cas, c'est souvent signe :

- d'une **modélisation non achevée** ou d'une redondance.

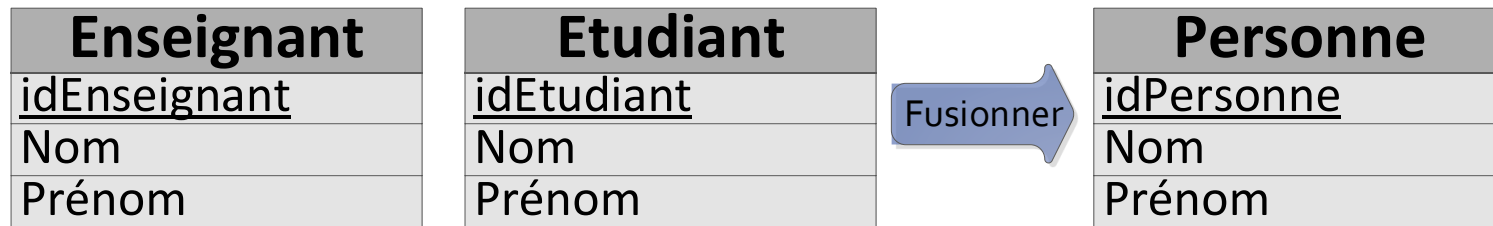
Ex: attribut « Adresse Facturation » présent dans un type entité « Facture » et dans un type entité « client »

- ou deux attributs ayant effectivement **un nom identique mais réellement distincts** (ambiguïté pour le système et pour l'utilisateur)

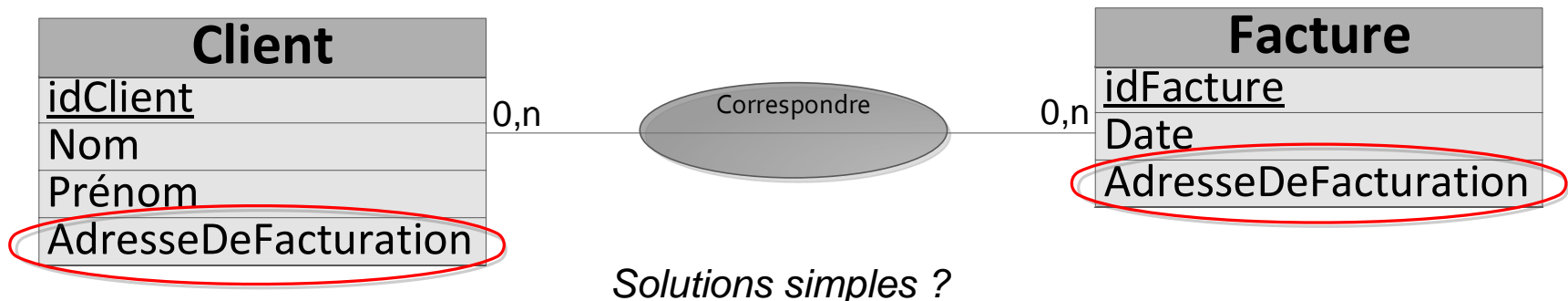
Ex: attribut « Nom » présent dans un type entité « Auteur » et « Abonné »

Factoriser et rationaliser

Modélisation non achevée



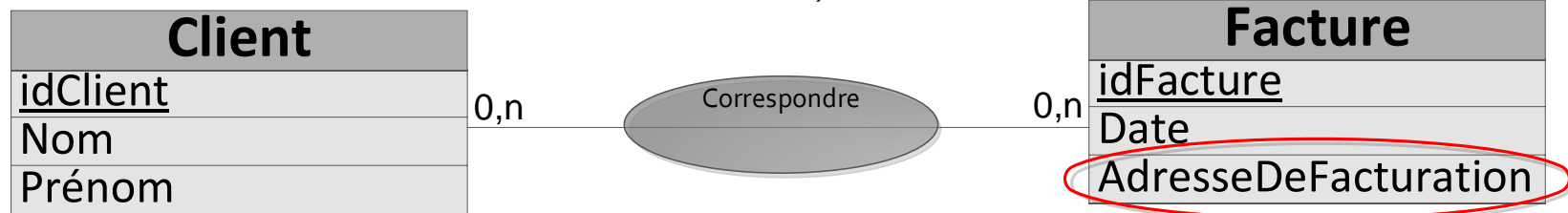
Redondance
(information identique)



Factoriser et rationaliser

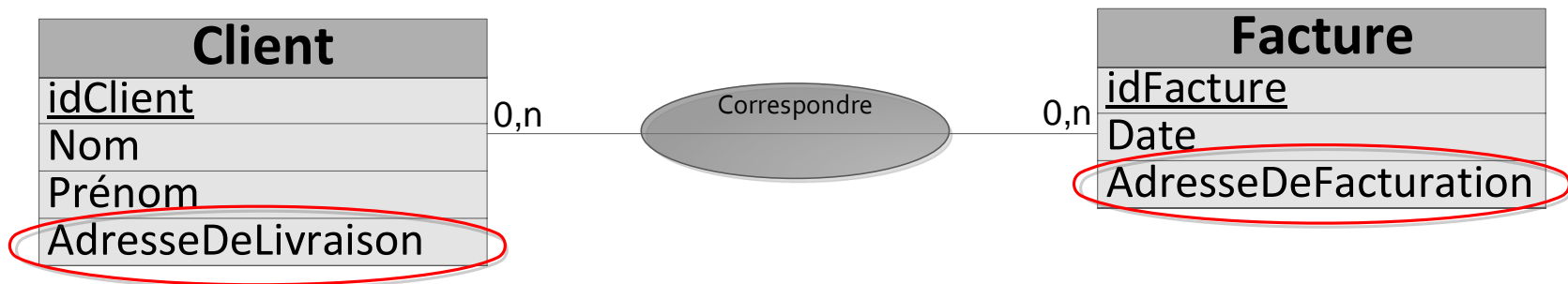
Rationaliser

(faire disparaître la redondance : sans ambiguïté et économie de place, une adresse validée à chaque facture)



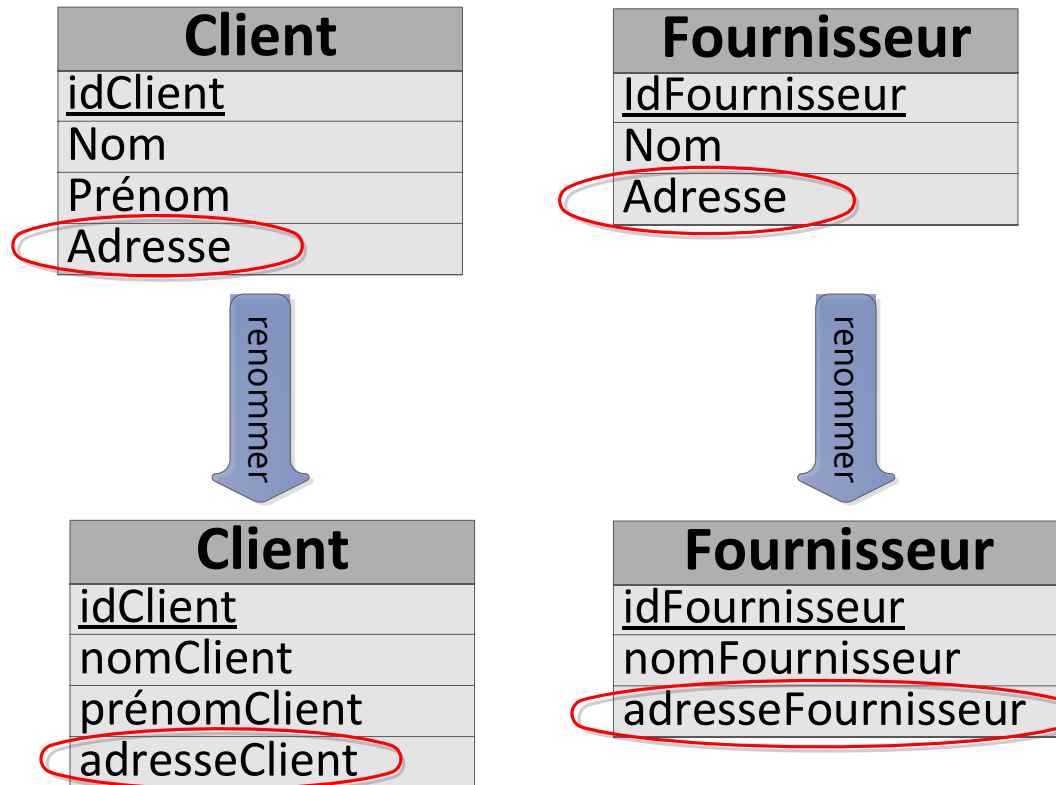
Renommer

(e.g. Amazon : deux adresses identiques ou différentes pour besoins fonctionnels)



Cas gérés par les SGBD...

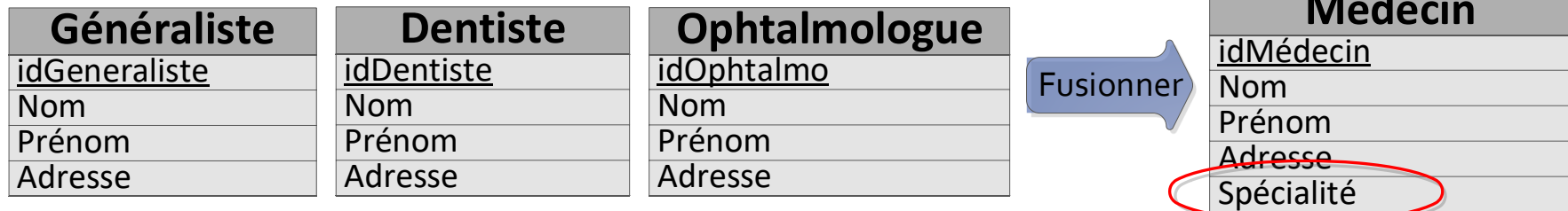
(noms identiques mais entités différentes)



...mais présente des risques de confusion lors de la réalisation des requêtes.

Quand fusionner ou supprimer des Type-Entités ou des Type-Associations ?

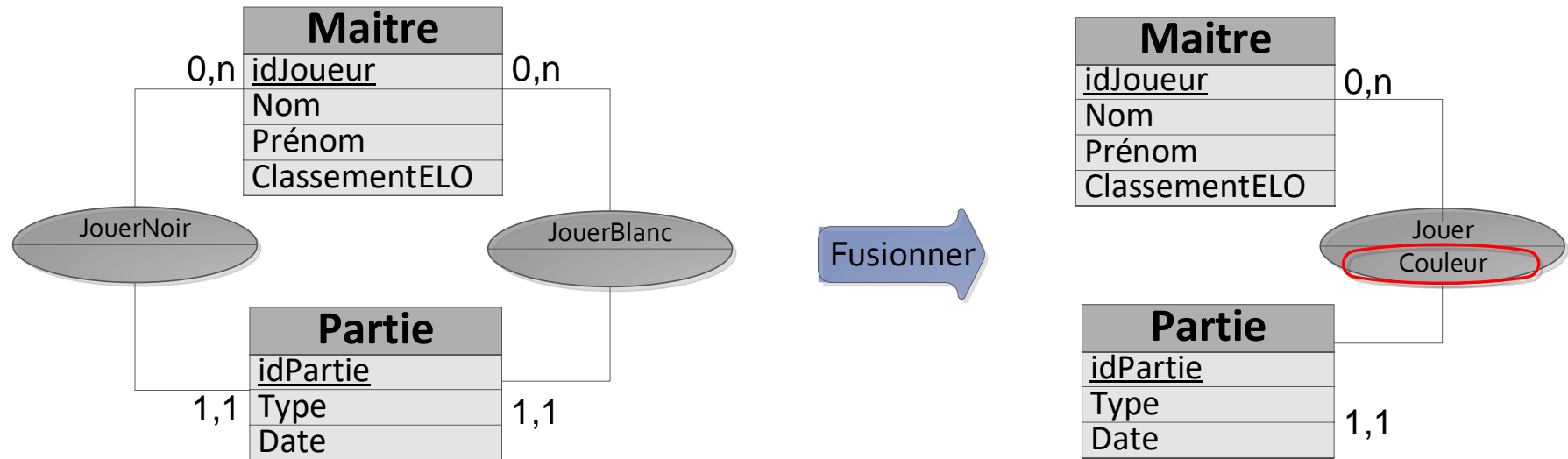
Règle 20 : Il faut **factoriser** les **Types-Entités** lorsque cela est possible



Mais pas toujours nécessaire : cela **dépendant des besoins** !

*Exemple précédent de la bibliothèque universitaire, où l'on cherche à avoir la capacité d'identifier les personnes qui sont à la fois auteurs et abonnés.
Dans ce cas la spécialisation offre les avantages des deux solutions !*

Règle 21 : Il faut factoriser les Type-Associations lorsque cela est possible



Ici, un seul Type-Association est nécessaire.

Type-Association de 1,1

Règle 22 : si toutes les cardinalités d'un Type-Association sont de 1,1 alors ce Type-Association n'a pas lieu d'être.

Fournisseur	
<u>idFournisseur</u>	1,1
nomFournisseur	
Adresse	

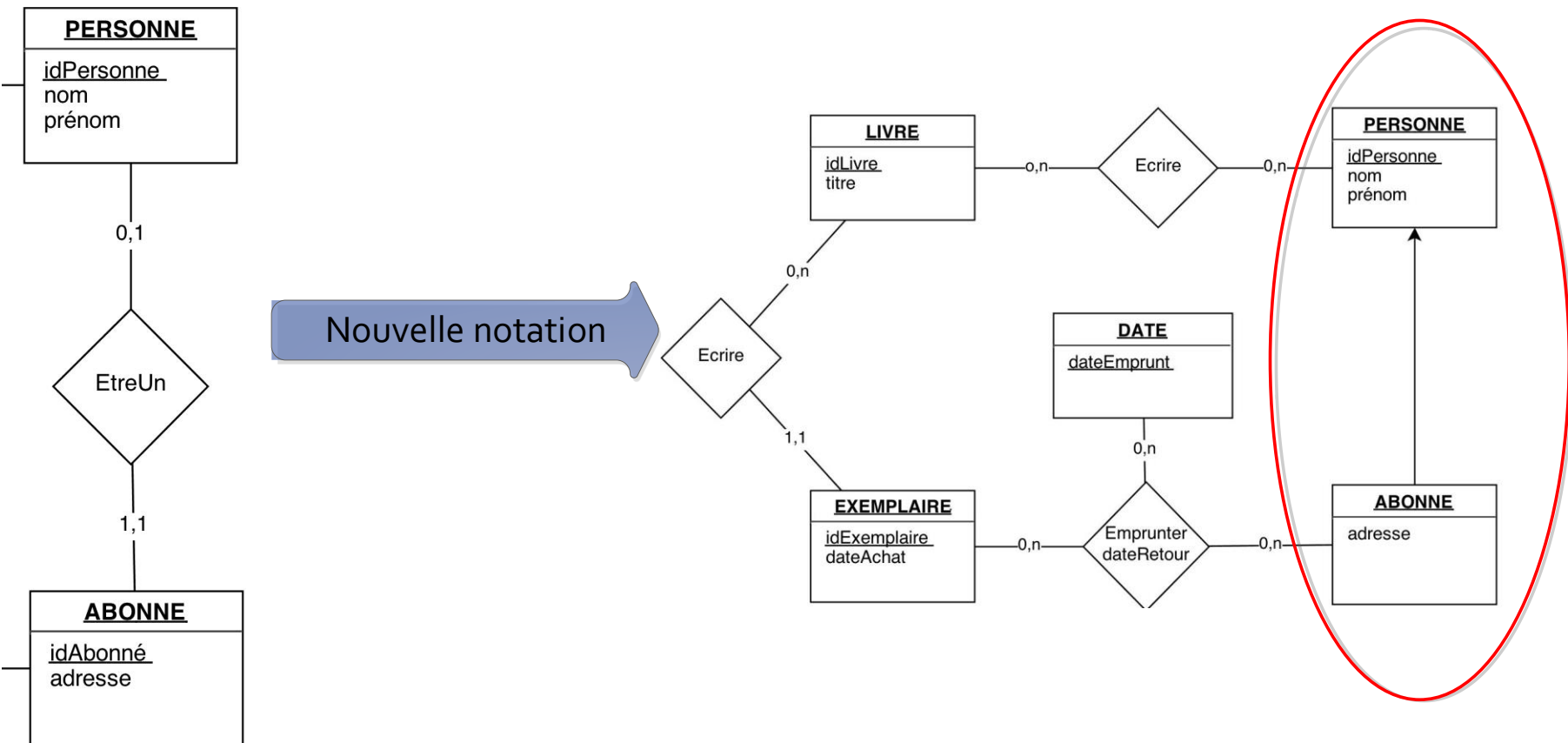


Fusionner

Fournisseur	
<u>idFournisseur</u>	
nomFournisseur	
Adresse	
nomContact	
numéroTéléphoneContact	

Type-Association de spécialisation

A l'inverse un Type-Association binaire possédant des **cardinalités de 0,1 et de 1,1** traduit généralement un **Type-Association de spécialisation**. Il ne doit généralement pas être supprimé.



Normalisation des Type-Entités et des Type-Associations

Les formes normales sont différents niveaux de qualité pour éviter :

- les **redondances**
- les **sources d'anomalies**

Nous parlerons des formes normales à deux reprises :

- au niveau du modèle **Entité-Associations** (approche plus intuitive)
- et au niveau du Modèle Relationnel (plus détaillé).

Chaque forme normale (n) vérifie les précédentes (n-1).

Rappel : la clé

Définition 12 : **Un identifiant, une clé**

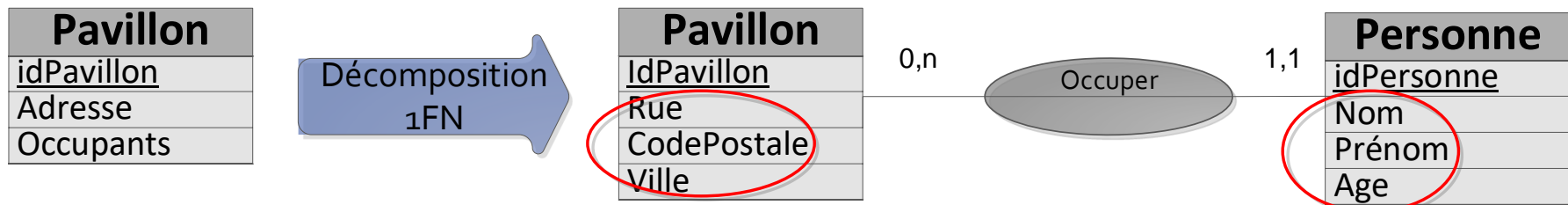
un identifiant (ou clé) d'un type-entité est un ensemble minimal d'attributs qui permet **d'identifier chaque entité de manière unique**.

Tout Type-Entité ou Type-Association possède au moins **un clé et peut en avoir plusieurs** :

- Dans ce cas, il y a **plusieurs clés candidates**;
- et l'identifiant choisi est nommée la **clé primaire**.

La première forme normale (1FN)

Définition 13 : (« **non décomposable** ») un Type-Entité ou un Type-Association est de première forme normale si, et seulement si, ses attributs sont non-décomposables.



La nature élémentaire des attributs dépend des **besoins exprimés** :

- une adresse peut être considérée comme élémentaire si elle est entière,
- mais si, par exemple, il est nécessaire de faire des statistiques sur les villes, elle devra alors être décomposée.

La deuxième forme normale (2FN)

Définition 14 : (« dépendance partielle ») un Type-Entité ou un Type-Association est de la deuxième forme normale si, et seulement si :

- il est de la première forme normale ;
- les attributs **ne dépendent pas d'un des attributs participant à la clé** (ou lorsque tous les attributs dépendent de la clé).

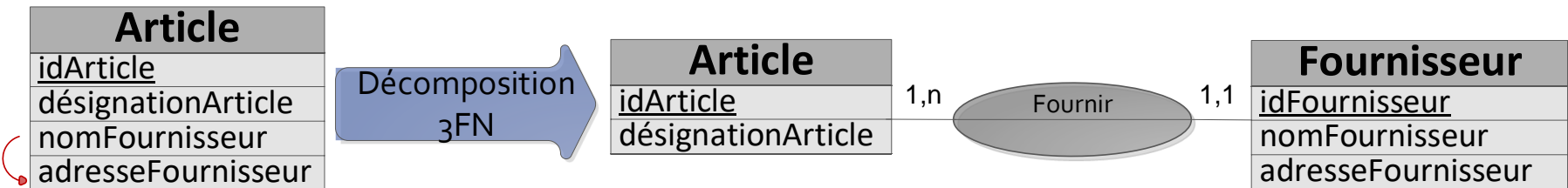


Dans cet exemple, un fournisseur peut avoir plusieurs articles et un article peut être fourni par plusieurs fournisseurs. **adresseFournisseur ne dépend que de nomFournisseur** (seulement une partie de la clé) → ajout clé arbitraire (mais dans cet exemple cela ne fait que déplacer le problème... 3FN)

La troisième forme normale (3FN)

Définition 15 : (« dépendance transitive ») un Type-Entité ou un Type-Association est de la troisième forme normale si, et seulement si :

- il est de la deuxième forme normale ;
- tout attribut n'appartenant pas à la clé **ne doit pas dépendre d'un ensemble d'attributs non-clé.**



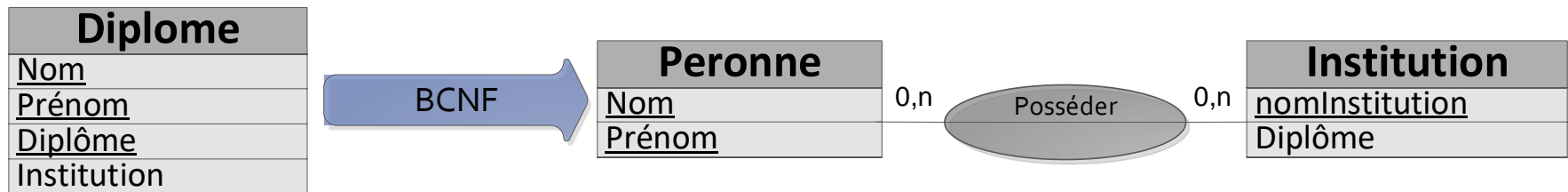
Dans cet exemple **adresseFournisseur** dépend de **nomFournisseur** (qui, lui, dépend de la clé).

➡ Il faut donc créer un nouveau Type-Entité Fournisseur

La forme normale de Boyce-Codd (BCNF)

Définition 16 : un Type-Entité ou un Type-Association est forme normale de Boyce-Codd si, et seulement si :

- il est de la troisième forme normale ;
- aucun attribut faisant partie de la clé dépend d'un **attribut ne faisant pas partie d'une des clés candidates**.

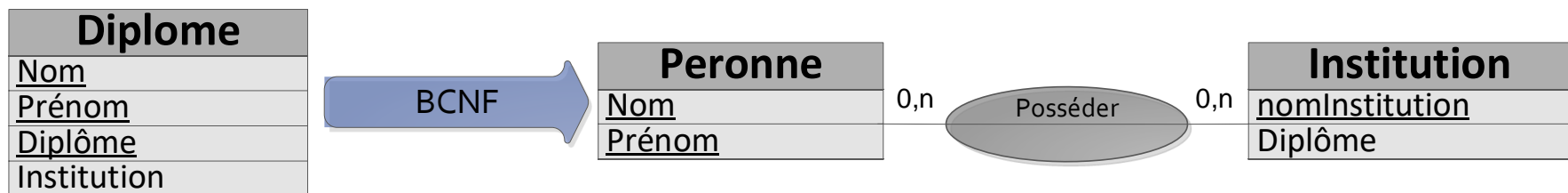


C'est donc une extension aux clés candidates de la troisième forme normale !

Dans cet exemple, il y a deux clés candidates :

- (nom, prénom, institution)
- (nom, prénom, diplôme)

Comme Diplôme dépend d'Institution, les deux clés candidates sont liées par une dépendance → décomposition du Type-Entité.



Notation simplifiée pour les Type-Associations de spécialisation

Notions avancées sur le modèle Entité-Association

- Bien choisir l'identifiant

- Des noms uniques

- Quand fusionner ou supprimer ?

- Normalisation des Type-Entités et des Type-Associations

 - Rappel : la clé

 - Première forme normale (1FN)

 - Deuxième forme normale (2FN)

 - Troisième forme normale (3FN)

 - Forme normale de Boyce-Codd (BCNF)

Synthèse : Etapes de conception d'un modèle Entité-Association

Introduction au modèle relationnel

- Introduction

- Définitions du Modèle Relationnel

 - Relation, Attribut et Domaine

 - Tuple

 - Degré et cardinalité

 - Les clés (candidate, primaire, étrangère)

 - Synthèse

 - Base de données et schéma de relation

 - Le schéma relationnel d'une base de données

- Passer du modèle Entité Association au Modèle Relationnel

 - Règles Générales

 - Cas particuliers des Type-Associations 1:1

 - Vidéo

Synthèse : étapes de conception d'un modèle Entité-Association

Il est possible d'identifier plusieurs étapes dans la conception du MCD (Modèle Conceptuel de Données).

Processus **non linéaire** et « **parallélisable** » :

- il est souvent nécessaire de revenir à l'étape précédente,
- et il n'est pas nécessaire d'avoir terminé une étape pour commencer l'étape suivante.

Les cinq étapes sont :

- Recueil des besoins
- Tri de l'information
- Identification des Type-Entités
- Identification des Type-Associations
- Vérification du Modèle

Synthèse : étapes de conception d'un modèle Entité-Association

1 // Recueil des besoins

- Etape essentielle visant à collecter les données à partir des documents d'entreprise et de tous support d'information.
- Un cahier des charges est un plus.

2 // Tri de l'information

- Etre attentif à la polysémie/synonymie : deux mots (ex : « référence de l'article » et « numéro de produit ») peuvent être utilisés pour caractériser la même chose.

3 // Identification des Type-Entités

- Repérer les attributs pouvant servir d'identifiant permet souvent de repérer un Type-Entité.
- Etre attentif au fait qu'un même concept du monde réel peut être dans un cas un attribut et dans un autre un Type-Entité (ex : Une marque de voiture peut être un attribut du Type-Entité Véhicule dans la base de données d'une préfecture, ou peut être considérée comme le Type-Entité Constructeur dans la base de données d'un magasin de location de voiture)

Synthèse : étapes de conception d'un modèle Entité-Association

4 // Identification des Type-Associations

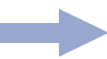
Repérer les Types-Associations (et leurs attributs éventuels) reliant les Type-Entités.

Il peut être difficile de faire un choix entre un Type-Entité et un Type-Association (étudier les cardinalités des TA retenus).

5 // Vérification du Modèle

Vérifier le modèle :

- en respectant les règles du cours 1 et 2
- et les formes normales;



Le Modèle Conceptuel de Données (MCD) ne s'intéresse pas aux **traitements futures** (calculs, statistiques...), toutefois il est important d'en avoir la connaissance pour préparer les **données élémentaires**.

Notation simplifiée pour les Type-Associations de spécialisation

Notions avancées sur le modèle Entité-Association

- Bien choisir l'identifiant

- Des noms uniques

- Quand fusionner ou supprimer ?

- Normalisation des Type-Entités et des Type-Associations

 - Rappel : la clé

 - Première forme normale (1FN)

 - Deuxième forme normale (2FN)

 - Troisième forme normale (3FN)

 - Forme normale de Boyce-Codd (BCNF)

Synthèse : Etapes de conception d'un modèle Entité-Association

Introduction au modèle relationnel

- Introduction

- Définitions du Modèle Relationnel

 - Relation, Attribut et Domaine

 - Tuple

 - Degré et cardinalité

 - Les clés (candidate, primaire, étrangère)

 - Synthèse

 - Base de données et schéma de relation

 - Le schéma relationnel d'une base de données

- Passer du modèle Entité Association au Modèle Relationnel

 - Règles Générales

 - Cas particuliers des Type-Associations 1:1

 - Vidéo

Le modèle relationnel

Proposé en 1970 par le mathématicien Edgar F. Codd, c'est un modèle simple basé sur la théorie des ensembles.

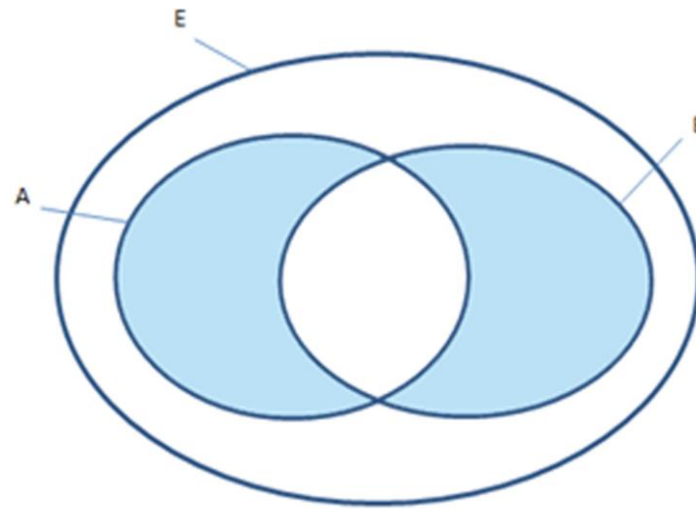


Diagramme de Venn

La manipulation d'ensembles avec le modèle relationnel correspond souvent à identifier la zone recherchée du diagramme de Venn.

(e.g. ensemble des étudiants ESIEE inscrits au bureau des sports et membres du bde)

Mais pas seulement la théorie des ensembles:
regroupements, somme, compte...

Le modèle relationnel poursuit toutefois deux objectifs :

- Garantir un degré d'indépendance avec la **représentation interne des données** (MPD, Modèle Physique de Données).
- Fournir une base solide pour traiter les problèmes de **cohérence** et de **redondance** de données.

Méthode Merise

Niveau Conceptuel :

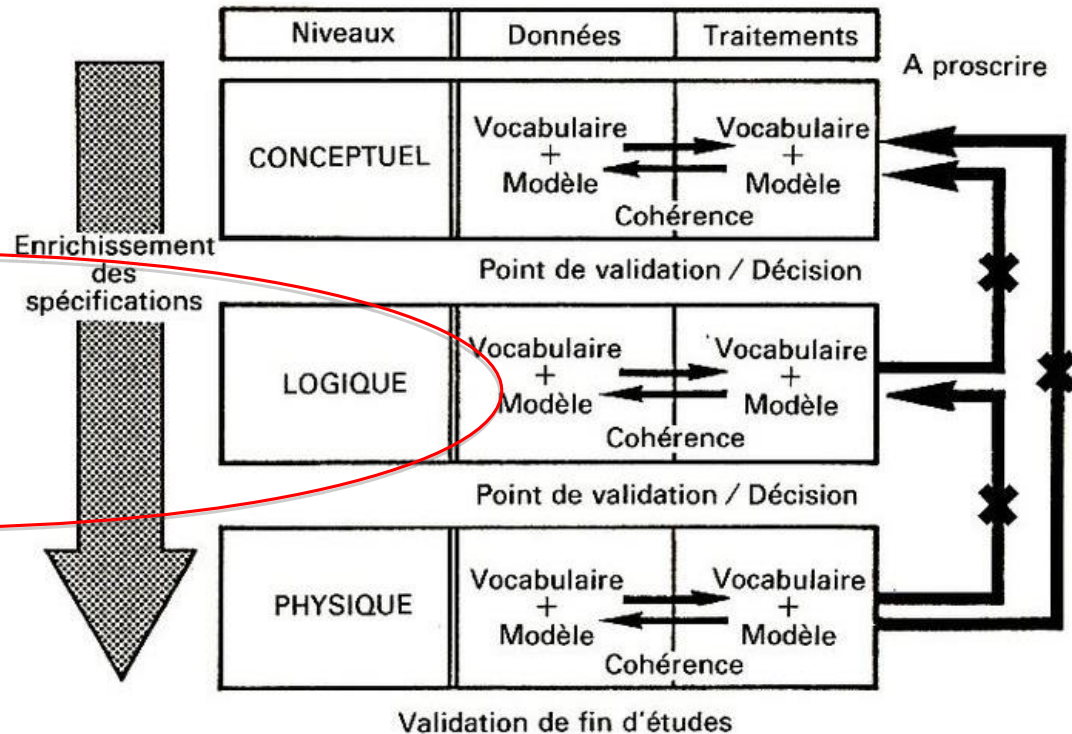
Produit le Modèle Conceptuel de Données (MCD)

Niveau Logique :

Produit le Modèle Logique de Données (le MLD est un MCD enrichi par la volumétrie, temps d'exécution...)

Niveau physique :

Produit le Modèle Physique de Données (description des supports et méthodes de stockages)



La méthode MERISE préconise d'analyser séparément données et traitements, à chaque niveau.

Merise (informatique), Wikipedia (01/01/2013), http://fr.wikipedia.org/wiki/Merise_%28informatique%29

MCD, Modèle Conceptuel de Données => Modèle Entité-Association

MLD, Modèle Logique de Données => Modèle Relationnel

Définitions du modèle relationnel

<i>PERSONNE</i>		
N° Sécu	Nom	Prénom
170022506862114	Durand	Caroline
182032632198742	Dubois	Jacques
146892301259421	Dupont	Lisa
184201236521455	Dubois	Rose-Marie

Définition 2.1 : la **relation** est une table avec des colonnes et des lignes.

Définition 2.2 : un **attribut** est une colonne nommée de la relation.

Définition 2.3 : un **domaine** est un ensemble de valeurs admissibles pour un ou plusieurs attributs.

PERSONNE		
N° Sécu	Nom	Prénom
170022506862114	Durand	Caroline
182032632198742	Dubois	Jacques
146892301259421	Dupont	Lisa
184201236521455	Dubois	Rose-Marie

Définition 2.4 : un **tuple** est une ligne dans une relation.

Termes formels	Alternative 1	Alternative 2
Relation	Table	Fichier
Tuple	Ligne	Enregistrement
Attribut	Colonne	Champ

<i>PERSONNE</i>		
N° Sécu	Nom	Prénom
170022506862114	Durand	Caroline
182032632198742	Dubois	Jacques
146892301259421	Dupont	Lisa
184201236521455	Dubois	Rose-Marie

Définition 2.5 : le **degré** d'une relation est le nombre d'attributs qu'elle contient. (ici, 3)

Définition 2.6 : la **cardinalité** d'une relation est le nombre de tuples qu'elle contient. (ici, 4)

Les clés

Définition 2.7 : une **clé candidate** d'une relation est un ensemble minimal des attributs d'une relation dont les valeurs identifient à coup sûr une occurrence

La valeur d'une clé candidate est donc distincte pour tous les tuples de la relation.

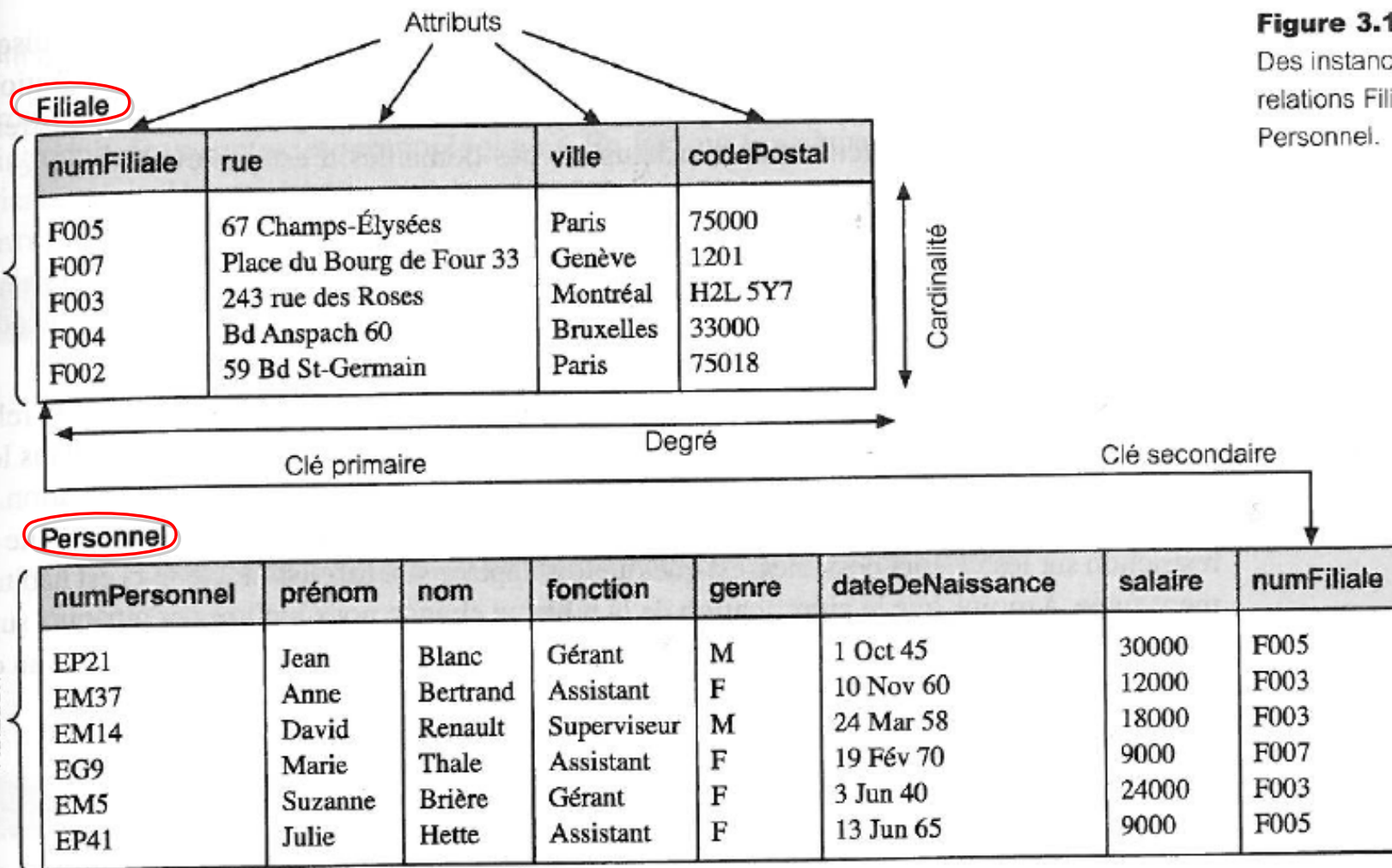
Définition 2.8 : une **clé primaire** d'une relation est une de ses clés candidates (soulignée).

Définition 2.9 : une **clé étrangère** dans une relation est formée d'un ou plusieurs attributs qui constituent une clé candidate dans une autre relation.

Recapitulatif

Figure 3.1

Des instances des relations Filiale et Personnel.



Définition 2.10 : une **base de données relationnelle** est une collection de relations normalisées, portant des noms de relations distincts.

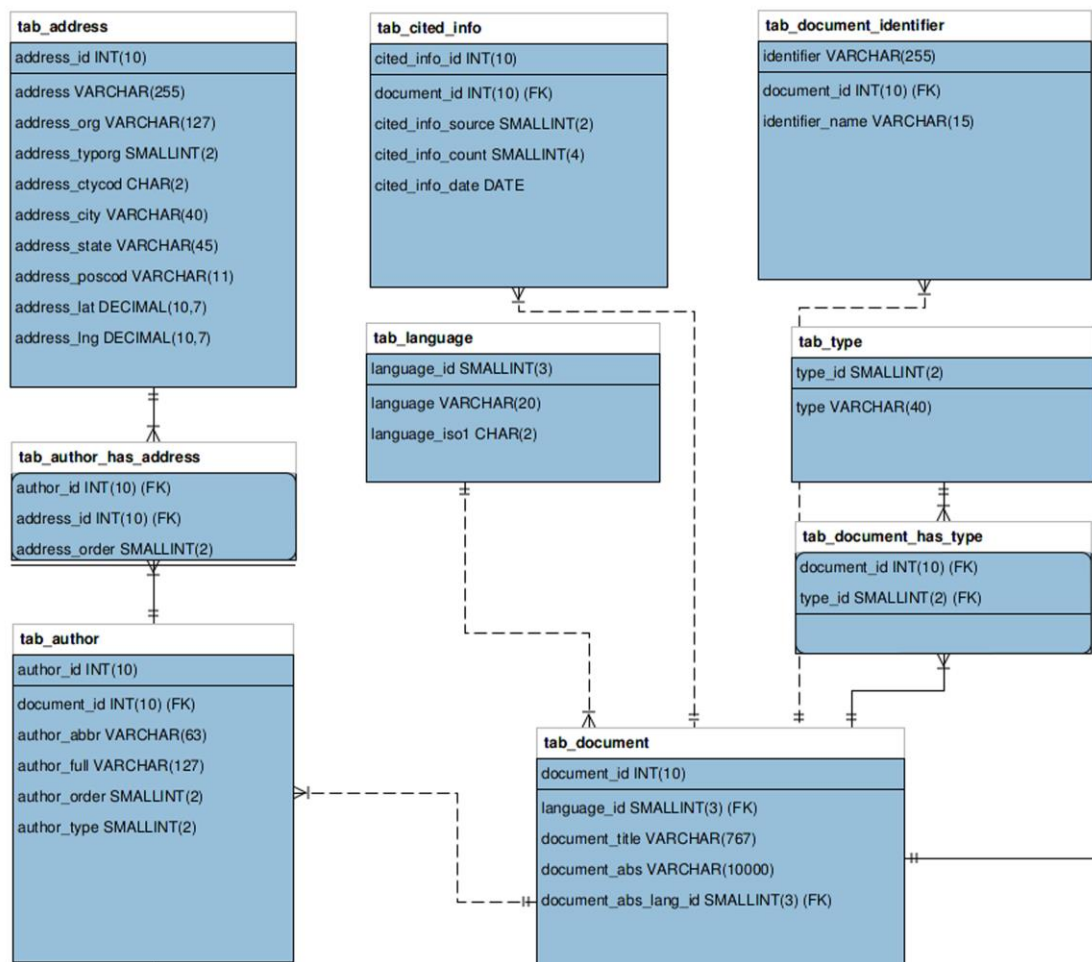
Définition 2.11 : le **schéma d'une relation** précise le nom de la relation ainsi que la liste des attributs avec leurs domaines.

PERSONNE		
N° Sécu	Nom	Prénom
170022506862114	Durand	Caroline
182032632198742	Dubois	Jacques
146892301259421	Dupont	Lisa
184201236521455	Dubois	Rose-Marie

Exemple de schéma de relation (attribut : domaine) :

PERSONNE(N° Sécu : Entier, Nom : Chaîne, Prénom : Chaîne)

Définition 2.12 : le **schéma relationnel** d'une base de données relationnelle est constitué par l'ensemble des schémas de relation.



Passer du modèle Entité-Association au Modèle Relationnel

Règle 2.13 : un Type-Entité du Modèle Conceptuel de Données (MCD) devient une **relation**, c'est à dire une table.

Règle 2.13a : son identifiant devient une **clé primaire** de la relation.

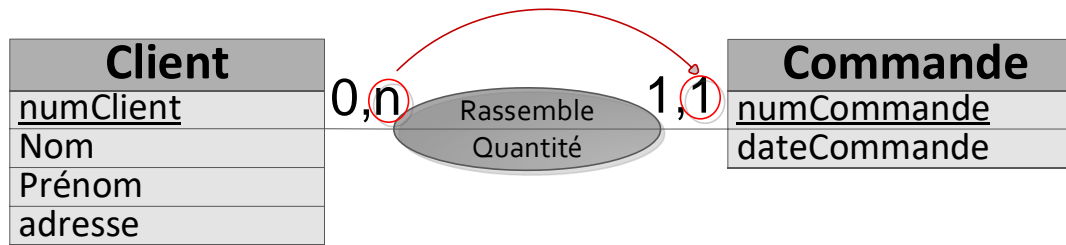
Règle 2.13b : les autres propriétés deviennent les **attributs** de la relation.

Client
<u>numClient</u>
Nom
Prénom
adresse

CLIENT(numClient, Nom, Prénom, adresse)

Type-Association de type "1:n"

Règle 2.14 : une Type-Association de type "1:n" se traduit par la création d'une **clé étrangère** dans la relation correspondante à l'entité du côté du "1"

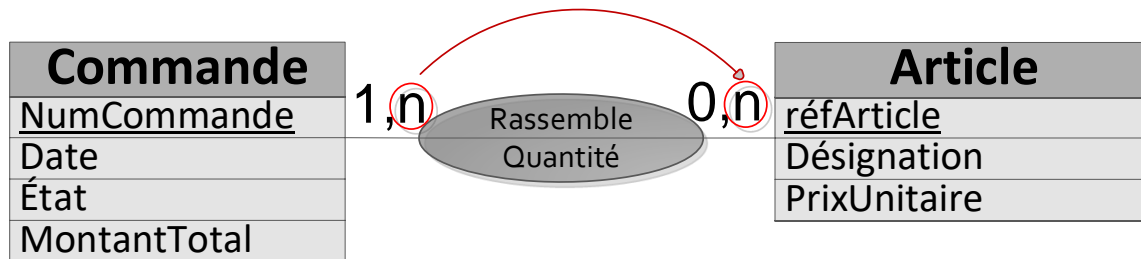


CLIENT(numClient , nom , prenom , adresse)

COMMANDE(numCommande, dateCommande, #numClient)

Type-Association de type "n:n"

Règle 2.15 : un Type-Association de type "n:n", ou de dimension >2 ou réflexif, devient une relation

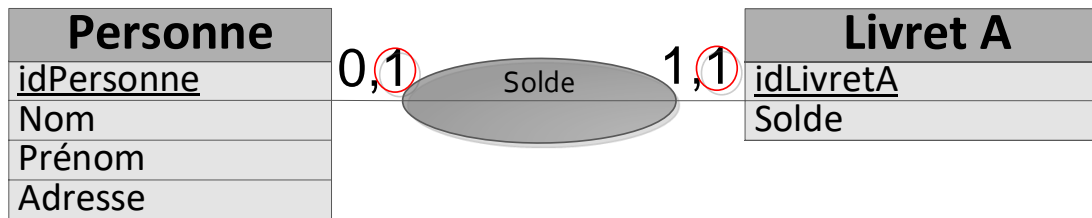


RASSEMBLE(#numero commande, #reference article, quantité)

ARTICLE(reference article, DésignationArticle, prixUnitaireArticle)

COMMANDE(numero commande, dateCommande, MontantCommande)

Cas particuliers des Type-Associations 1:1



La règle générale nous autorise à mettre une clé étrangère dans les deux relations. La cardinalité 0,1 de Personne montre qu'une personne peut ou pas avoir un livret A, ce qui nous permet de choisir :

Livret A(idLivretA, solde, #idPersonne)

Personne(idPersonne, nom, prénom, adresse)