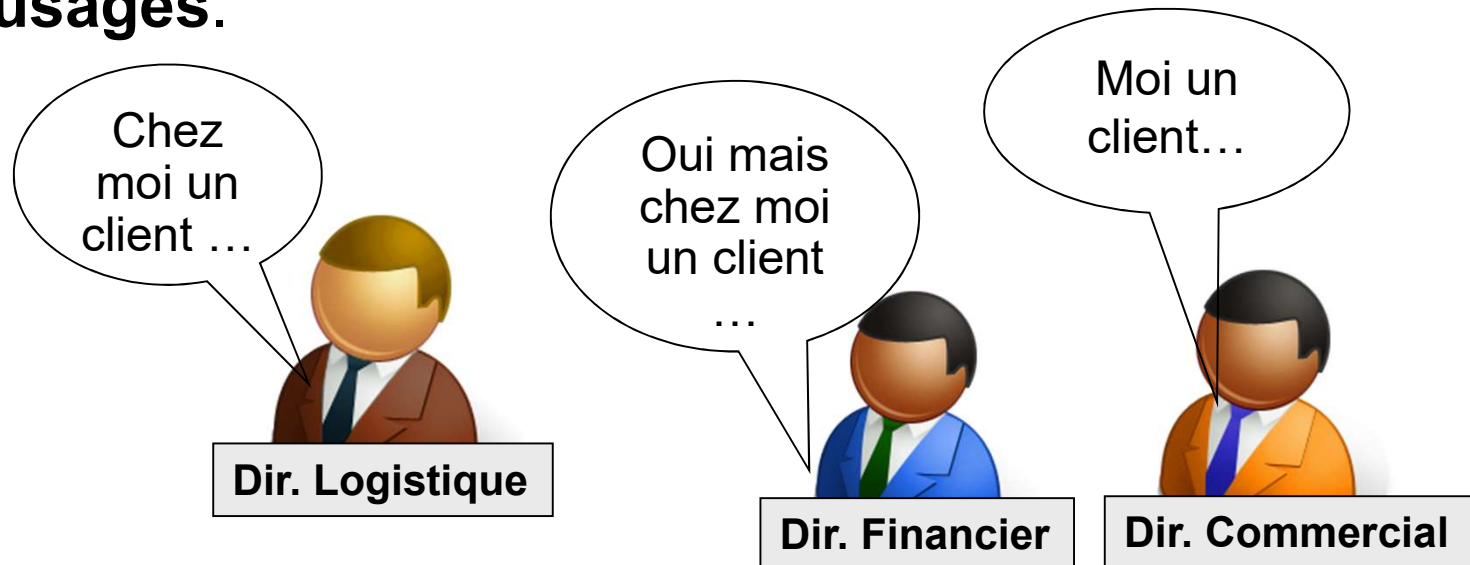


# **Le modèle conceptuel des données MCD**

# La problématique des données

- Il ne suffit pas de s'intéresser au nom et aux propriétés des données : type, longueur, valeurs.
- Il faut s'intéresser à la donnée elle-même, ses **sens** et ses **usages**.



**Les acteurs** peuvent utiliser les mêmes mots avec des sens ou des contenus différents (synonymes, polysèmes).

## Exemple réel : sens du mot « client »

Libellé FR	Libellé EN	Sens
<b>Client</b>	<b>Customer</b>	Correspond à l' <b>adresse principale d' un donneur d' ordres</b> depuis laquelle on reçoit les ordres de réalisation des prestations. Exemple de « client » : Kraft Foods France
<b>Client opérationnel</b>	<b>Operational customer</b>	Est la déclinaison d' un « <b>client</b> » pour un <b>lieu géographique ou un métier particulier</b> Exemple de « Client Opérationnel » : Kraft entreposage CPN
<b>Client de facturation</b>	<b>Bill-to customer</b>	Désigne le <b>tiers destinataire des factures d' un « client opérationnel »</b> .
<b>Client payeur</b>	<b>Payer customer</b>	Désigne le <b>tiers qui paye les factures d' un « client opérationnel »</b> .
<b>Client de gestion</b>	<b>Controlling customer</b>	Désigne <b>un ou plusieurs « clients opérationnels » dont les coûts et les recettes sont regroupées</b> . Le « client de gestion » est une notion propre aux contrôleurs de gestion.

Il faut comprendre les données ... avant de les décrire  
(dictionnaire des données).

Il faut aussi se poser des questions sur la **qualité** des données existantes et les exigences de qualité.

n° ss	Nom	Prénom	Date naissance	sexe	Adresse	Code Postal	Ville	Téléphone
171046734543621	Dupond	Albert	10/04/1971	F	3, rue de la gare	99999	Strasbourg	01 32145678
268065415498494	Durant	Lise	18/06/1968	F	Rue des Lilas	54000	Nancy	0345762345
268065415498494	Durant	Lisa	18.06.1968	F		54000	Null	0345762345

**Contradiction**  
(pb de cohérence)

**Hors nomenclature**  
(pb de conformité et d'intégrité)

**Incohérence**

**Doublon**  
(pb d'unicité)

**Erreur de saisie**  
(pb d'exactitude)

**Erreur de Format**  
(pb de conformité)

**Absence de valeur**  
(pb de complétude)

**Pb d'intégrité référentielle**

# Objectif du MCD

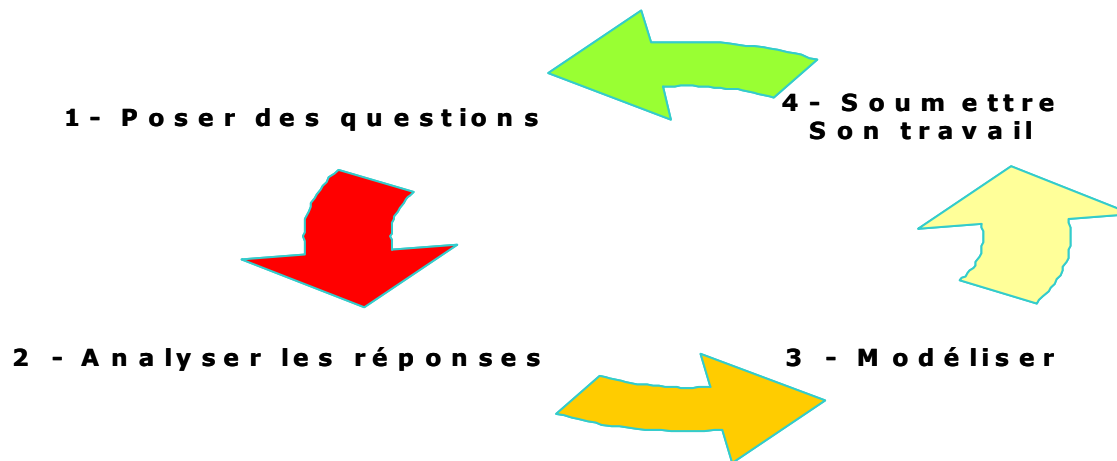
Décrire les données du SI, indépendamment de tout choix d'implantation physique.

## 1. Le dictionnaire des données

- **Inventaire** des données du domaine étudié.

Questions :

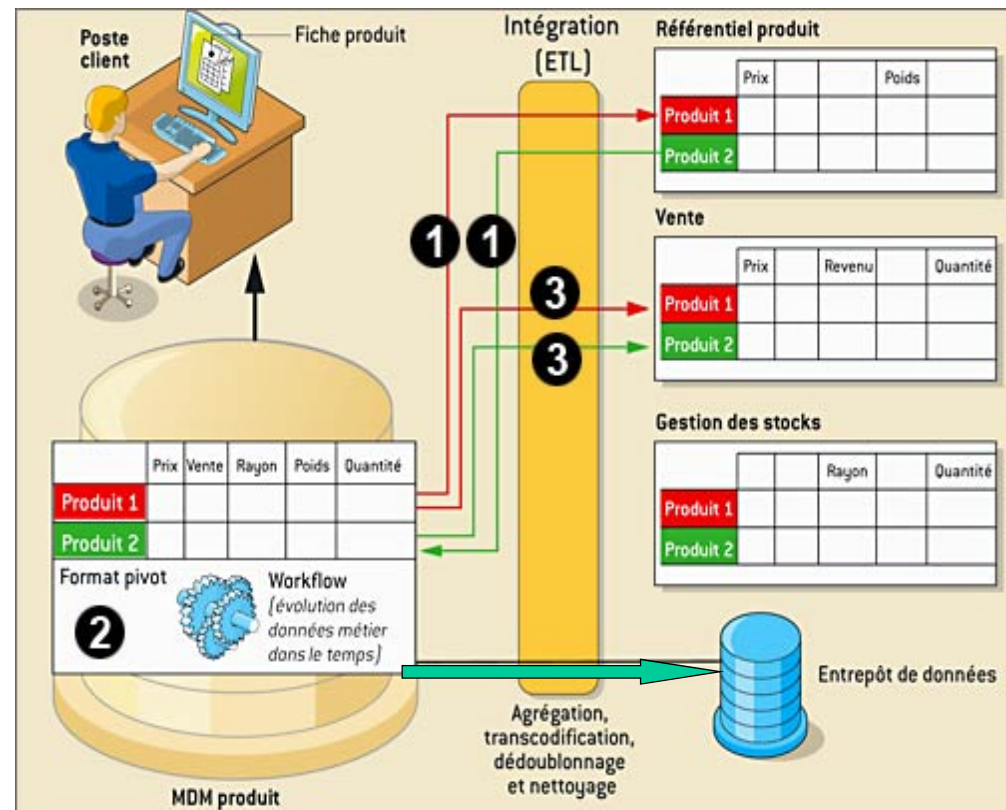
- sens pour les différents interlocuteurs; les différents sens sont à conserver.
- exigences de qualité et caractéristiques.



# Métadonnées (dictionnaire de données)

## Principe

Collecter et concentrer toutes les données relatives à un objet métier (client, fournisseur, produit...) dans un **format pivot**, puis transférer les mises à jour vers l'entrepôt de données.



- 1 Création ou mise à jour d'un objet métier
- 2 Conversion au format pivot
- 3 Synchronisation des informations

## Nombreuses caractéristiques :

- identificateur (mnémonique),
  - description (« sens » précis),
  - type (numérique, alphanumérique, ...),
  - taille,
  - mode d'obtention :
    - donnée mémorisée,
    - donnée calculée,
    - donnée "paramètre" : donnée utile à un traitement et non mémorisée (ex : date d'édition),
  - règle de calcul (pour les données calculées),
  - contraintes d'intégrité : intervalle de valeurs, liste de valeurs...
  - origine (document, système, service)
  - **volume,**
  - **fréquence des mises à jour,**
  - etc.
- } aspects quantitatifs

Descriptif **très simplifié** utilisé dans les exercices où toutes ces caractéristiques ne sont pas toujours disponibles :

documents

Rubrique	Description	Type	Mode	D1	D2	D3	D4
identificateur	libellé	entier	mémorisée		x		x
		réel	calculée	x			
		date	paramètre	x	x	x	x
		chaîne					
		booléen					



## **2. Le modèle conceptuel des données : le modèle entité/association**

- a) Concepts de base du modèle E/A.**
- b) Vérification et normalisation du modèle E/A.**
- c) Contraintes d'intégrité du modèle E/A**

## a) Les concepts de base

**Entité** : tout objet concret ou abstrait ayant une existence propre et conforme aux besoins de gestion de l'organisation.

Ex : le client «Dupond», le produit de référence «a456»...

**Classe d'entités (ou entité-type)** : ensemble des entités décrites par les mêmes caractéristiques.

Ex : la classe CLIENT dont «Dupond» est une occurrence (ou instance).

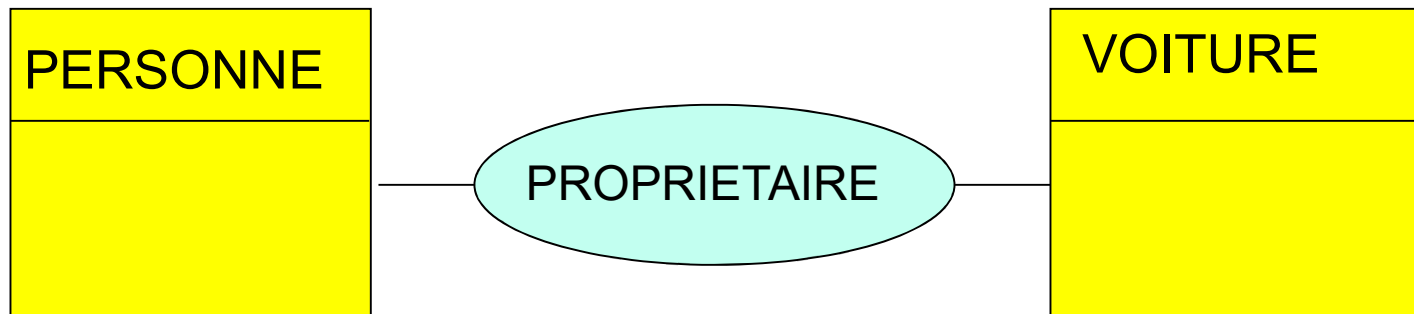
**Association** : n-uplet d'entités « sémantiquement liées ».

Ex: («Dupond», «1367 VS 54») indiquant que la personne Dupond est propriétaire de la voiture immatriculée 1367 VS 54.

**Classe d'associations (ou association-type) :** regroupe toutes les associations constituées des mêmes types d'entités jouant le même rôle dans l'association.

Ex: PROPRIETAIRE(PERSONNE, VOITURE)

Les occurrences de cette classe d'association sont un sous ensemble du produit cartésien PERSONNE x VOITURE (c.à.d. une partie de l'ensemble des couples possibles de personnes et de voitures).



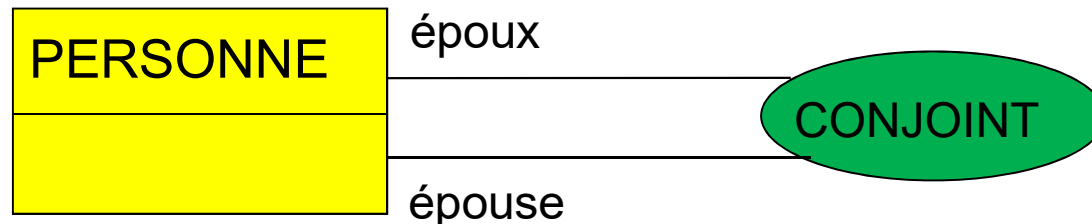
## Remarques

- On peut avoir plusieurs classes d'associations sur les mêmes classes d'entités.

Ex : PROPRIETAIRE(PERSONNE, VOITURE)  
et CONDUIRE(PERSONNE, VOITURE)

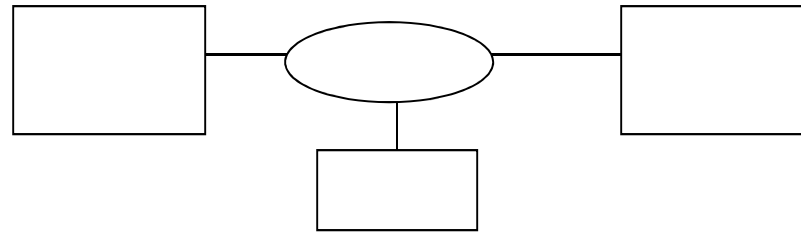
- On peut avoir une classe d'association sur une seule classe d'entités (on parle d'association 'réflexive'). On ajoute souvent dans ce cas des noms de **rôles** pour distinguer les deux occurrences.

Ex : CONJOINT(PERSONNE, PERSONNE)



- On peut avoir une classe d'association définie sur n classes d'entités (association n-aire n ou de dimension n ou à « n pattes »).

Ex: COURS(MATIERE, CLASSE, PROF)



Attention : les n pattes sont rares. Elle dénotent souvent des faiblesses dans l'analyse.

Relation à 2 : 80%

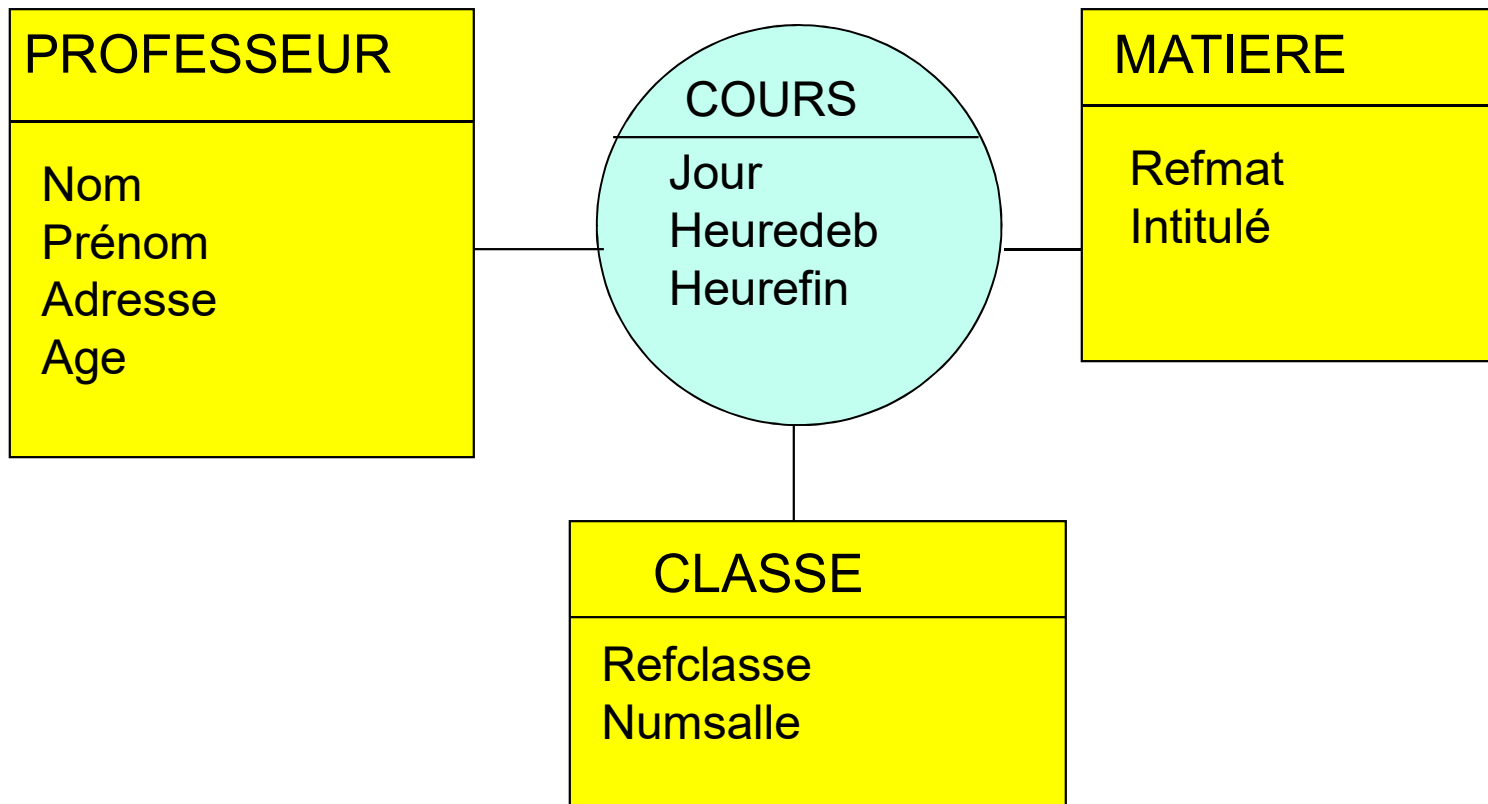
Relation à 3 : <20%

Relation à > 3 : .5 %

**Propriété** : donnée élémentaire permettant de caractériser les entités et associations

Ex : Nom, Prénom propriétés de PROFESSEUR

Jour, Heuredéb propriétés de COURS



**Identifiant** : propriété ou groupe de propriétés permettant d'identifier de manière unique chaque occurrence de la classe d'entités.

Ex : *N° immatriculation* pour VOITURE. Nom ne suffit pas pour PERSONNE. *N° Client* pour CLIENT (propriété ajoutée)

Les identifiants sont en général soulignés.

**Cardinalités** : indiquent pour chaque classe d'entités de la classe d'association, les nombres mini et maxi d'occurrences de l'association pouvant exister pour une occurrence de l'entité.

**La cardinalité minimum est 0 ou 1.**

**La cardinalité maximum est 1 ou n.**

**Une cardinalité** minimum à 0 signifie qu'il est possible d'observer (un jour) une occurrence d'entité sans occurrence d'association.

Donc 4 combinaisons possibles :

{	0,1	au plus 1
	1,1	1 et 1 seul
	1,n	au moins 1
	0,n	un nombre quelconque

Ex: PROPRIETAIRE(PERSONNE [0,n], VOITURE [1,1])

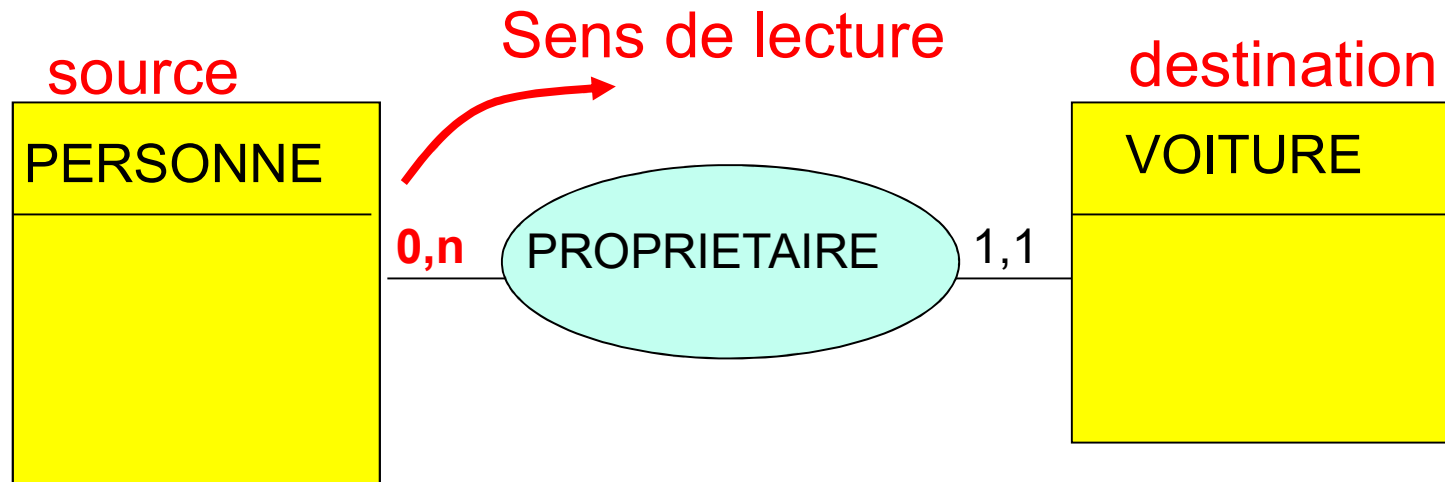
Une personne a 0 à n voitures; une voiture a 1 et 1 seul propriétaire.



CONDUIT(PERSONNE [0,n], VOITURE [1,n])

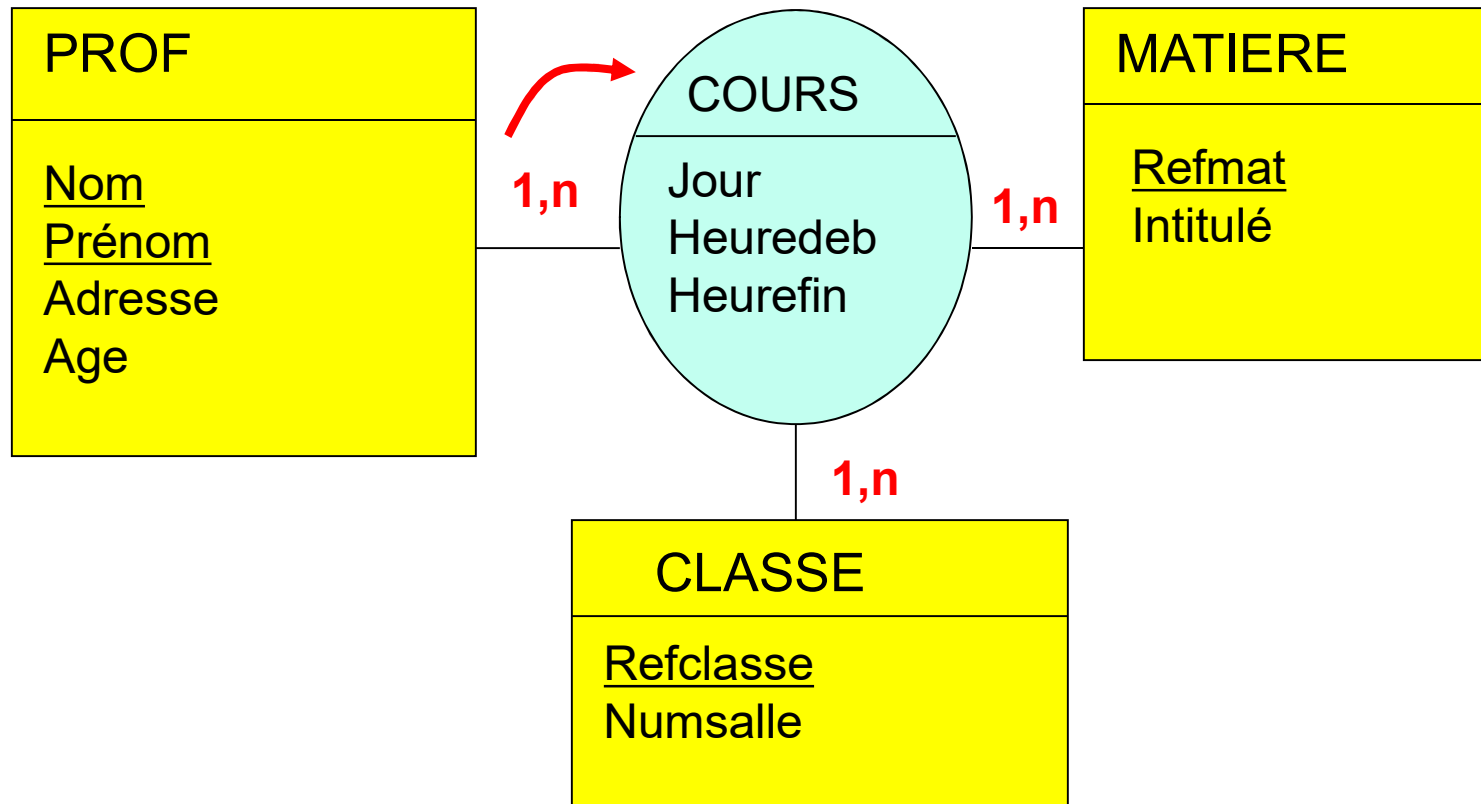
Une personne conduit 0 à n voitures; une voiture est conduite par 1 à n personnes.

Représentation graphique :



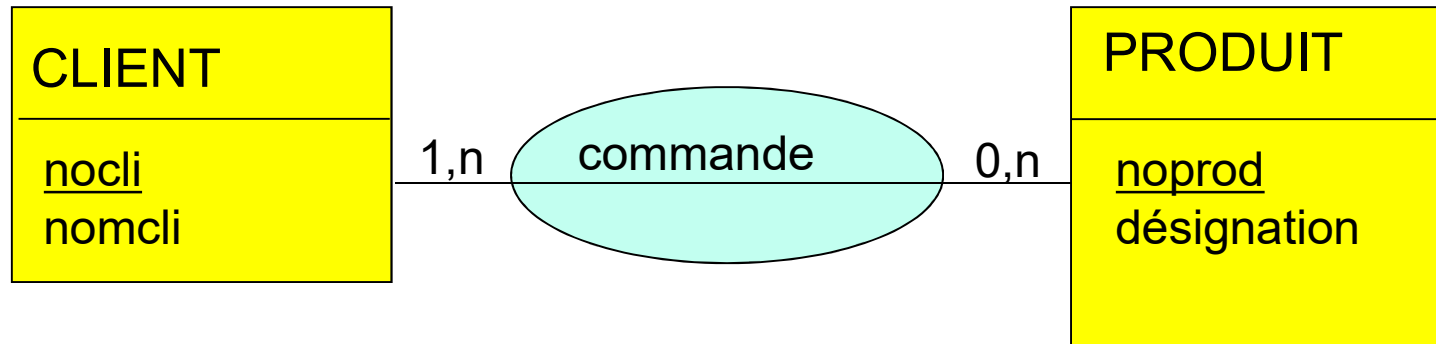
COURS(MATIERE [1,n], CLASSE [1,n], PROF[1,n])

Un prof. a 1 à n cours dans la semaine, une matière a 1 à n cours dans la semaine, une classe a 1 à n cours dans la semaine.



# Difficultés : choix entre entité et association ?

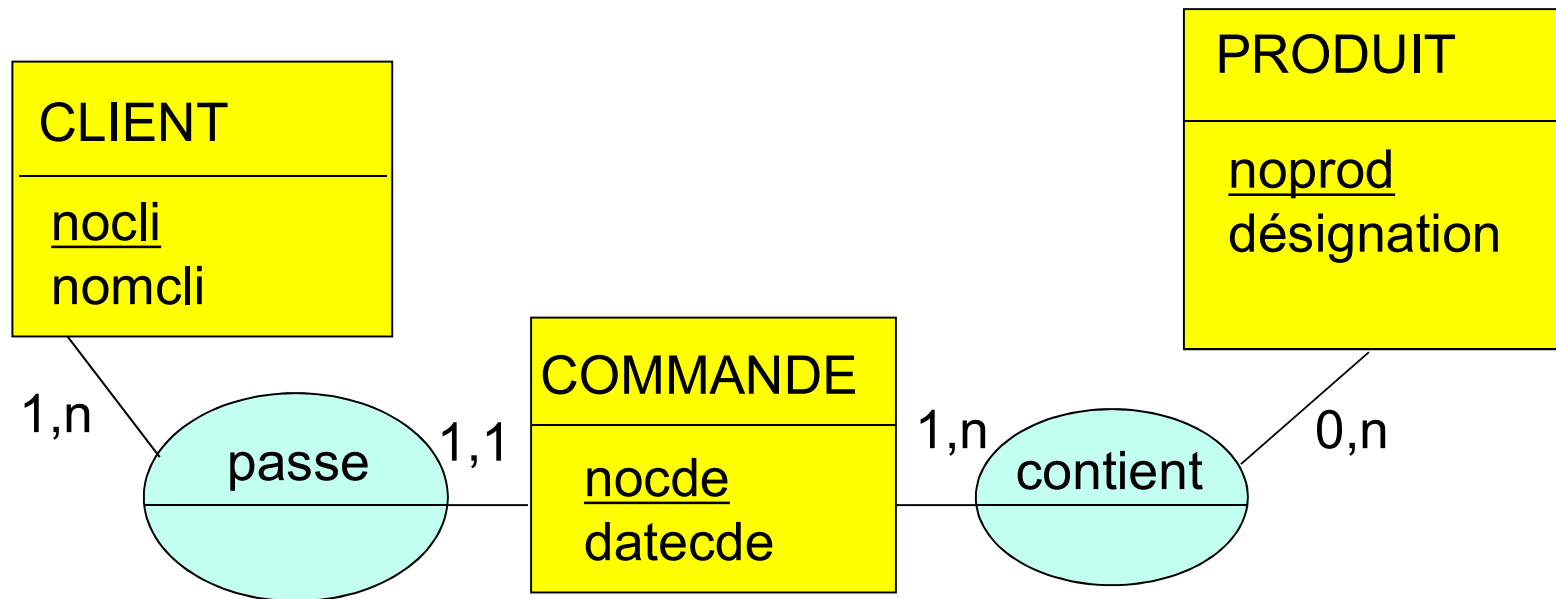
## 1) Solution avec association



Dans cette première solution la commande n'est pas une entité gérée pour elle-même. Elle existe tant que le client et le produit existent.

Ce peut être le SI du domaine 'fabrication' : on a juste besoin de savoir que les produits sont destinés à des clients.

## 2) Solution avec entité



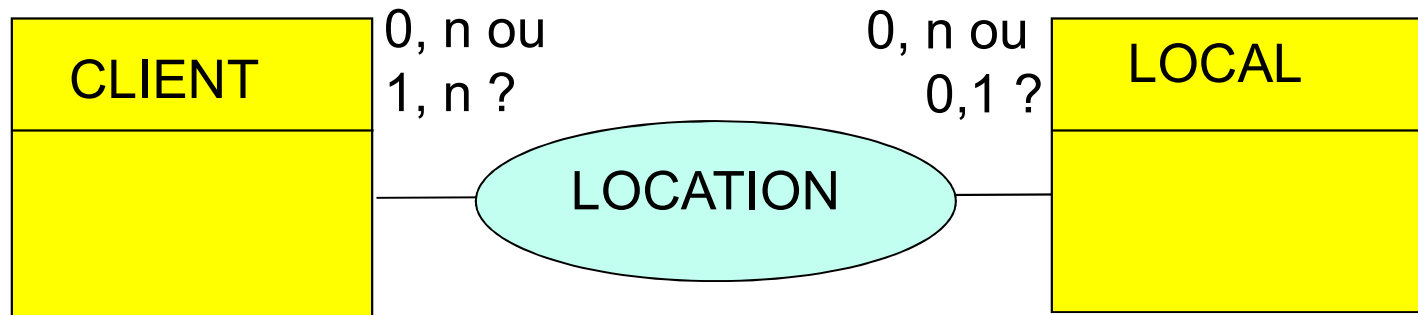
Dans cette seconde solution, les commandes sont identifiées (identifiant nocde) et décrites : on les gère en tant que telles. Elles peuvent être conservées même si le produit ou le client n'existent plus.

Ce peut être le SI du domaine financier.

## Quelques « critères » de choix

- Une entité a une **existence propre** et un **identifiant**.
- Une association n' existe que si ses extrémités existent et n' a pas d' identifiant explicite.
- Une entité peut être associée à d' autres entités, une association non.

## Difficultés : choix des cardinalités ?



Un **client** peut il avoir 0 location ? Est-ce encore un client ?

Un **local** peut il être loué plusieurs fois ? Non si la base représente une situation instantanée et si le local n'est pas partageable. Oui si on gère un historique ou si le local est partageable.

**Les cardinalités sont élément essentiel pour définir la sémantique (signification) des données, pas une « décoration » accessoire. Derrière cette notion on trouvera des contrôles (par le SGBD ou les programmes).**

Pour une situation donnée, **il n'existe pas une «solution» unique.**

Le « bon modèle » est celui qui est **accepté par les personnes concernées par le projet.**

## b) Vérification et Normalisation

Contrôler la qualité du modèle vis-à-vis :

- des fondements du modèle d'une part (règles de vérification),
- de la redondance de données d'autre part (règles de normalisation) .

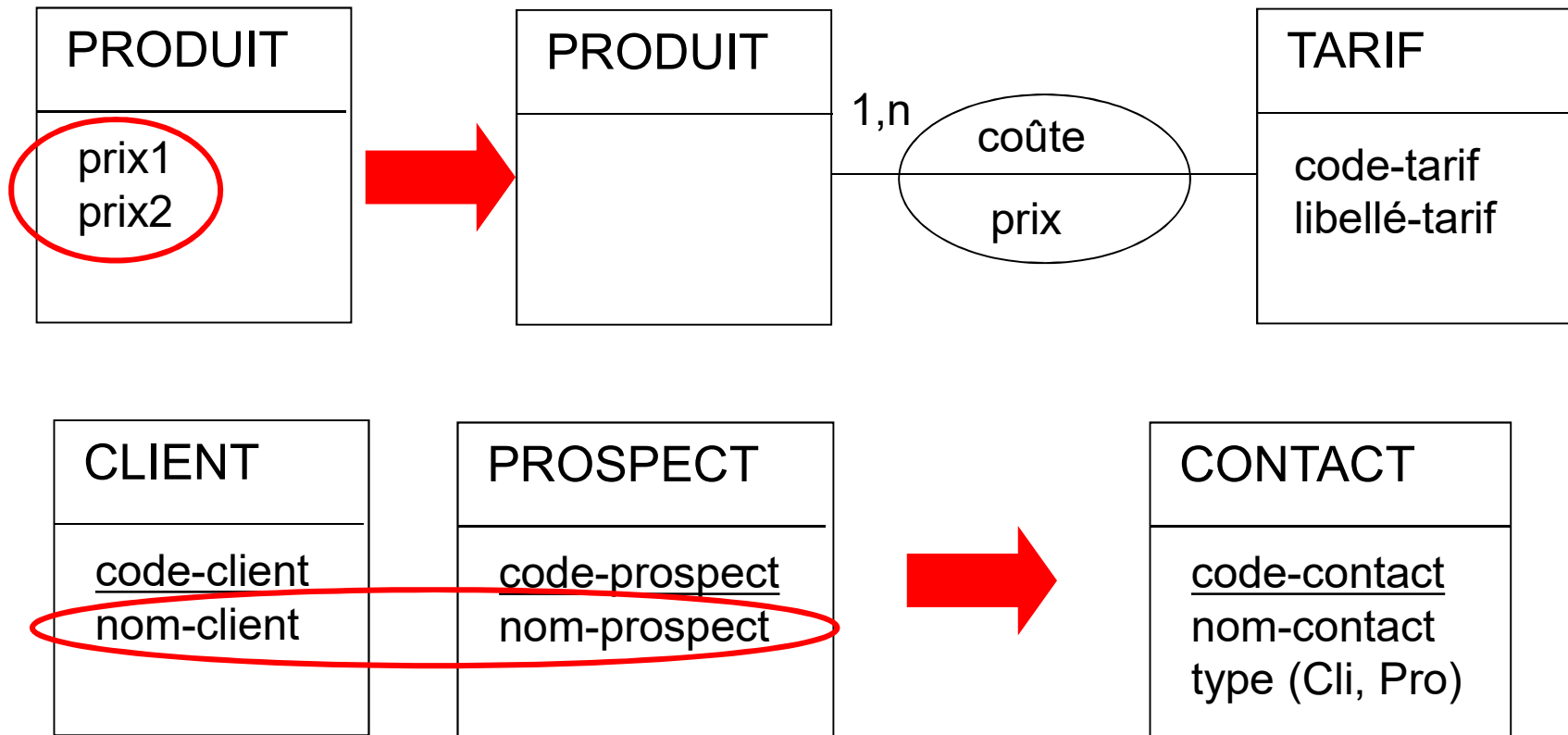
Permet de détecter certaines incohérences dans la construction des modèles.

### 1. Règles Générales

- Toute propriété doit apparaître une seule fois dans un modèle.

Il faut éliminer la **redondance** des propriétés dans la même entité (avec des noms différents) ou dans des entités distinctes :





Pas d'héritage dans le modèle E/A de base !

- Toutes les propriétés identifiées doivent apparaître dans le modèle.

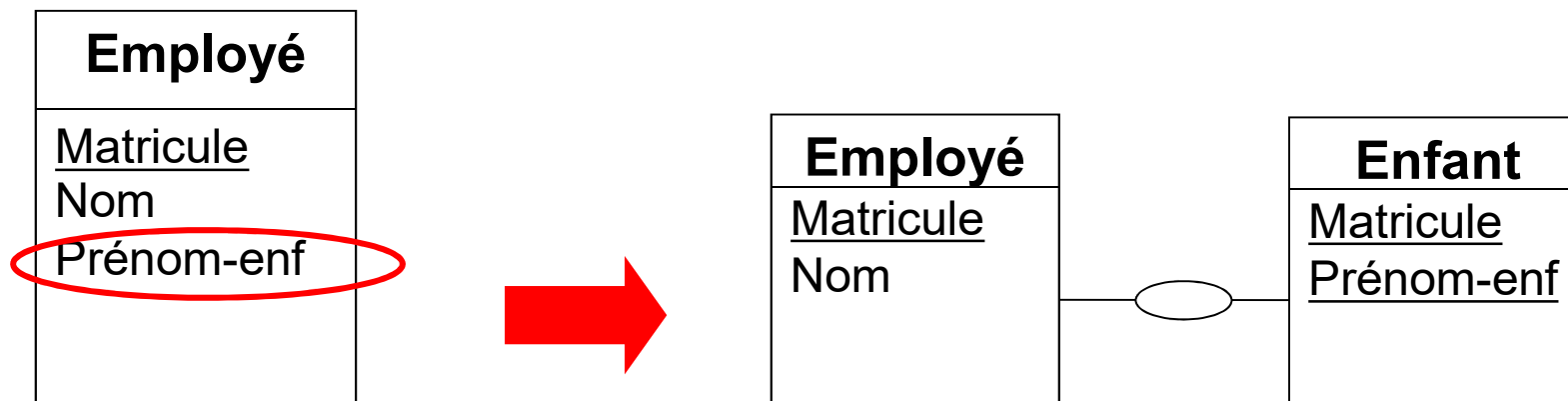
## 2. Règles sur les entités

### 2.a Règle de l'identifiant

Toutes les entités ont un identifiant.

### 2.b Règle de vérification des entités

Pour une occurrence d'une entité, chaque propriété ne prend qu'une seule valeur

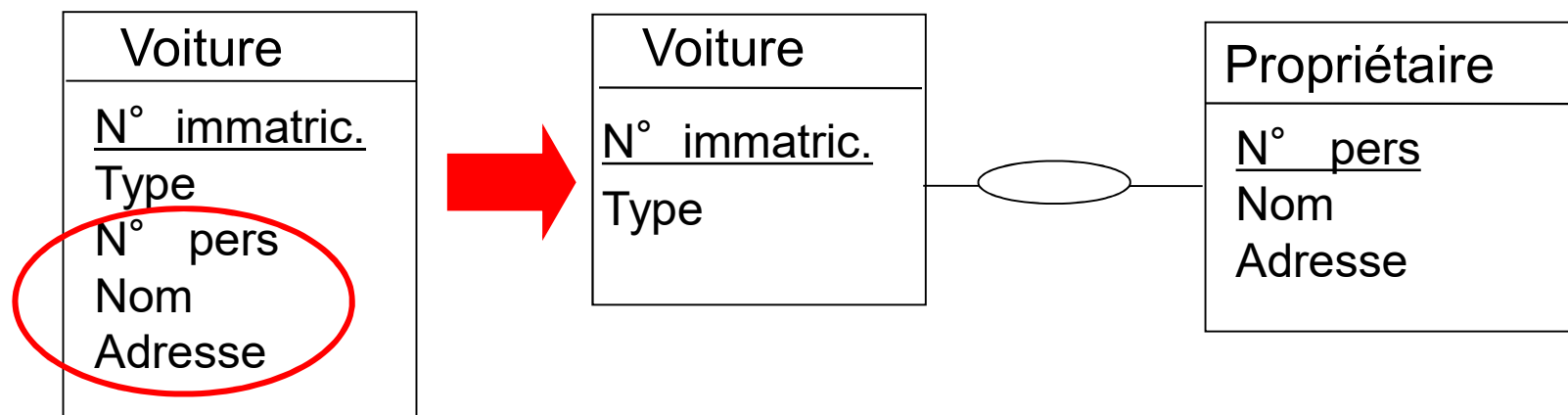


On décompose l'entité Employé en deux entités : Employé, et Enfant

## 2.c Règles de normalisation des entités

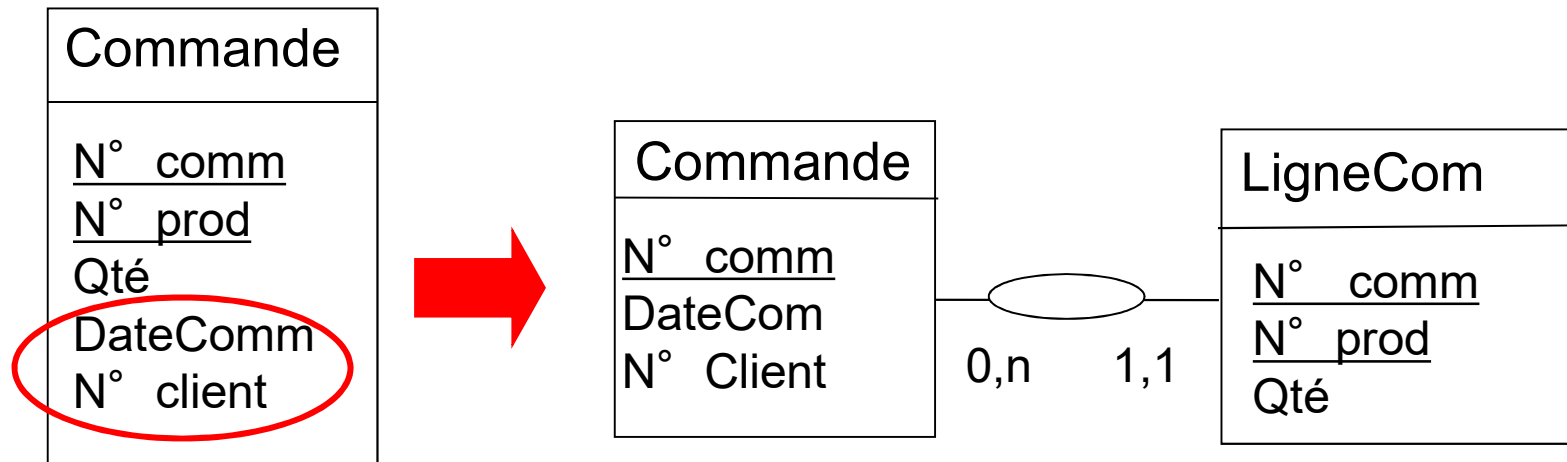
a) Les dépendances fonctionnelles (DF) entre les propriétés d'une entité doivent vérifier la règle suivante : **toutes** les propriétés de l'entité dépendent fonctionnellement de l'identifiant **et uniquement** de l'identifiant.

Rappel :  $\exists$  une DF  $X \rightarrow Y$  si à une valeur de  $X$  correspond une et une seule valeur de  $Y$  (réciproque pas vraie).



La DF:  $N^{\circ} \text{ pers} \rightarrow \text{Nom, Adresse}$  contredit la règle.

b) Une partie de l'identifiant ne peut pas déterminer certaines propriétés.



La DF  $n^{\circ} - comm \rightarrow date-comm, n^{\circ} - client$  contredit la règle. On décompose l'entité Commande en deux entités.

Ces règles correspondent aux 2FN et 3FN du modèle Relationnel (dépendance pleine et directe des clés).

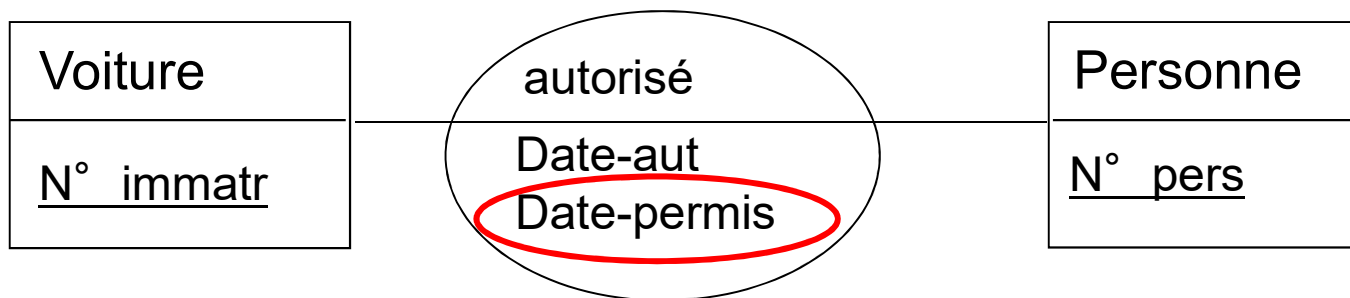
### 3. Règles sur les associations

#### 3.a Règle de vérification des associations

Pour une occurrence d'association, chaque propriété ne prend qu'une seule valeur.

#### 3.b Règle de normalisation sur les propriétés des associations

Toutes les propriétés de l'association doivent dépendre fonctionnellement de tous les identifiants des entités portant l'association, et uniquement d'eux.

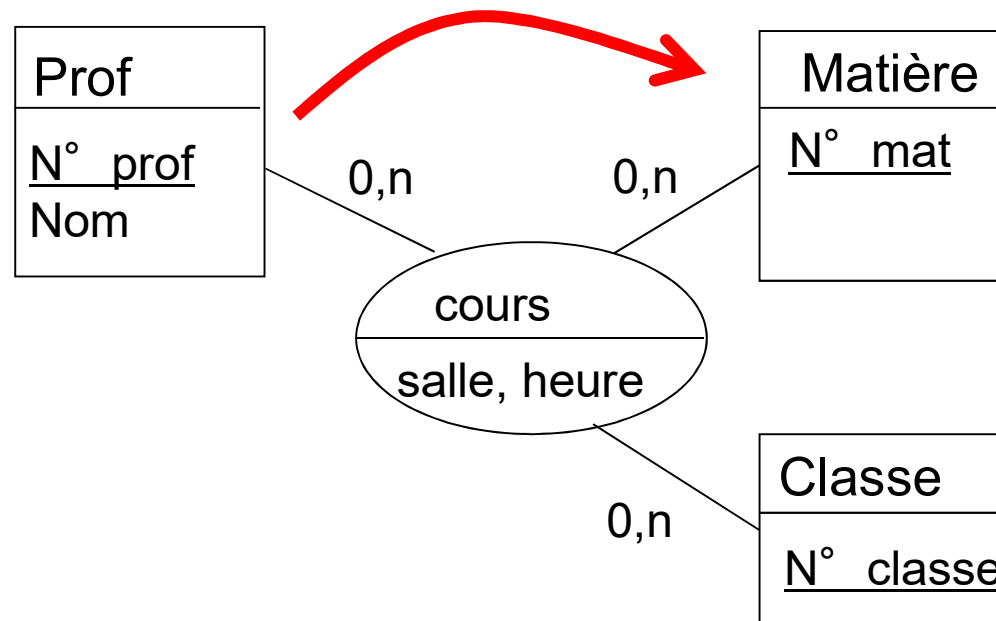


$N^{\circ} - pers \rightarrow Date-permis$  pose problème (donc déplacer Date-permis vers Personne)

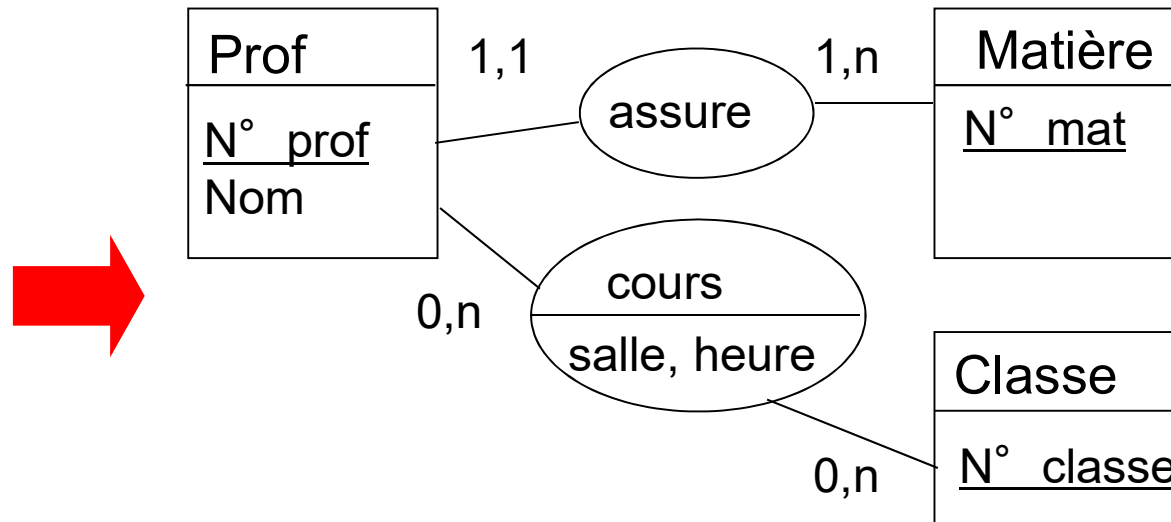
### 3.c La décomposition des associations n-aires

Il faut garder un minimum d'associations d'arité 2.

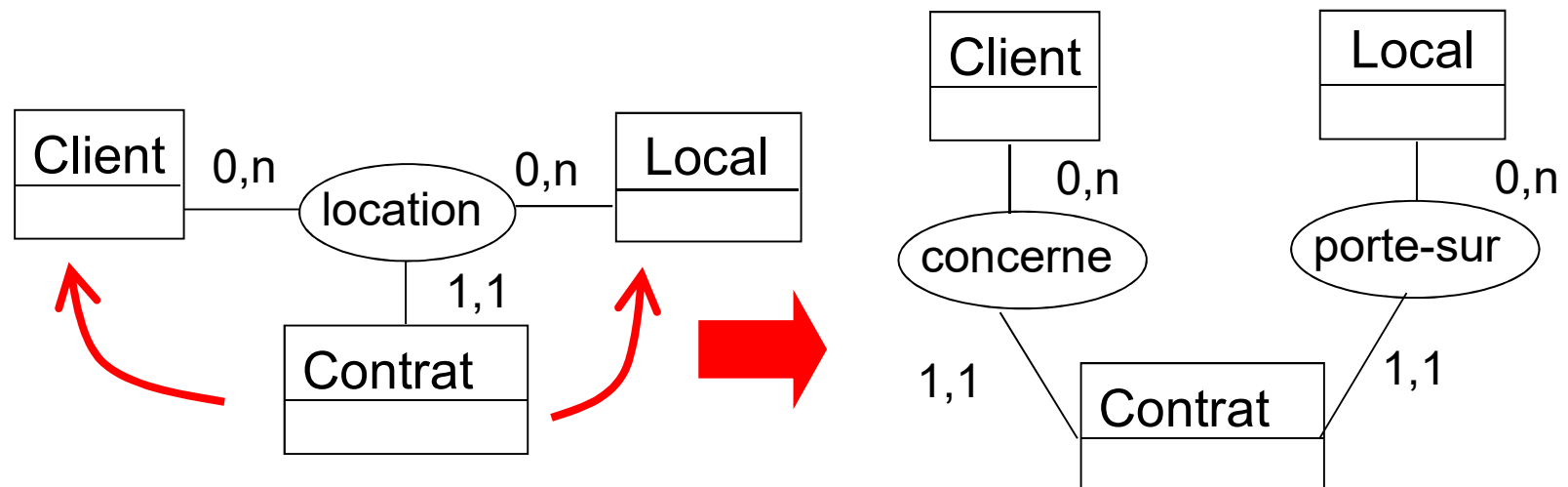
Si on observe une DF entre deux identifiants, on peut décomposer l'association n-aire.



Une éventuelle DF  $N^{\circ} \text{ prof} \rightarrow N^{\circ} \text{ mat}$  (c.à.d. si un prof enseigne une seule matière) conduit à la décomposition :

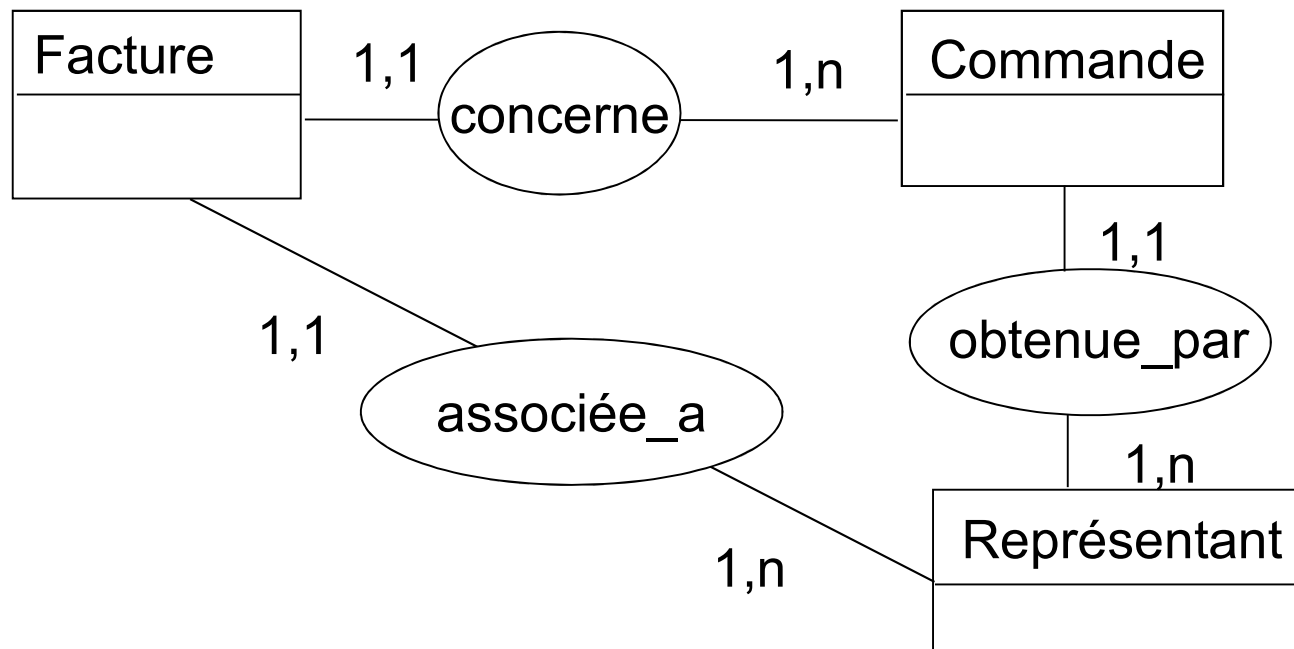


C' est le cas, quand une patte a une cardinalité 1,1.  
 Par exemple à 1 contrat est associé un client et un local :



### 3.d La suppression des associations transitives

Toute association pouvant être obtenue par transitivité de  $n$  autres associations peut être supprimée. La transitivité s'évalue en fonction de la **signification** des associations.



On supprime l'association *associée\_a*, car elle peut être obtenue par transitivité sur les associations *concerne* et *obtenue\_par*



## c) Quelques contraintes d'intégrité importantes

Les CI définissent des propriétés qui doivent être vérifiées par les données de la base.

### 1. Contraintes intégrées au modèle E/A

#### 1.a *Contrainte d'identifiant*

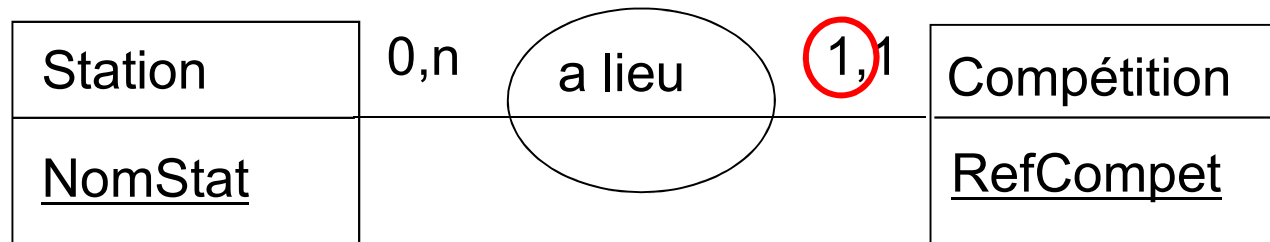
Les valeurs prises par l'identifiant sont uniques (dans le temps) et toujours définies.

Ex : identifiant de l'entité PERSONNE

- nom + prénom pas suffisant
- n° téléphone pas stable dans le temps
- n° SS réglementé (autorisation de la CNIL car danger de rapprochement de fichiers)

#### 1.b *Contraintes de cardinalité*

Les cardinalités portées par les entités membres d'association imposent des nombres minis et maxis d'occurrence dans l'association.



**Une cardinalité mini de 1 rend l'existence d'une occurrence d'entité dépendante de l'existence d'une occurrence d'une autre entité.**

Une compétition ne peut exister que si la station où elle se déroule existe.

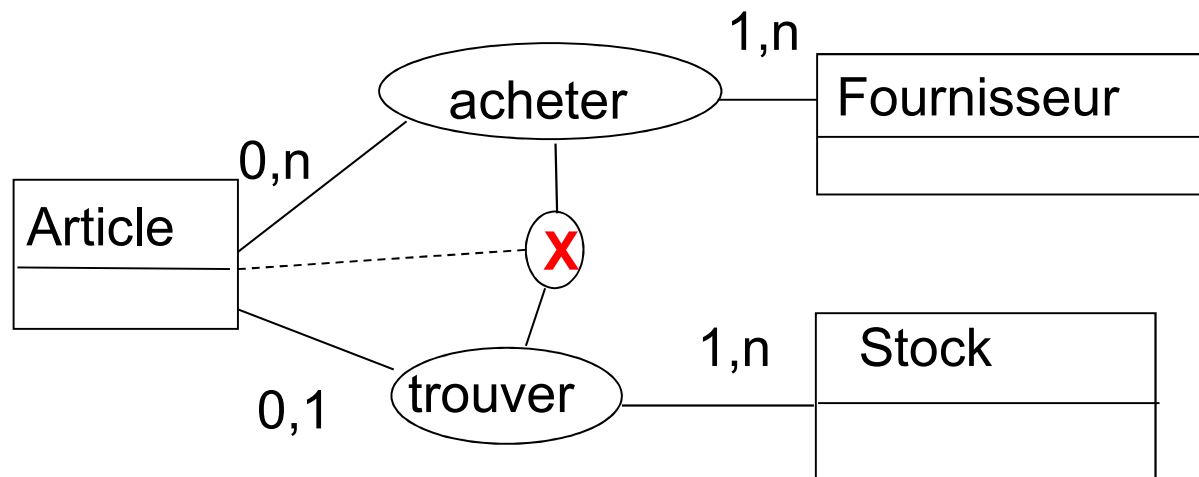
Une station peut exister de manière indépendante de toute compétition.

## 2. Contraintes extensions du modèle E/A

**Exemple : contraintes de participation des entités aux associations.**

### *2.a Exclusivité de participation d'une entité à plusieurs associations*

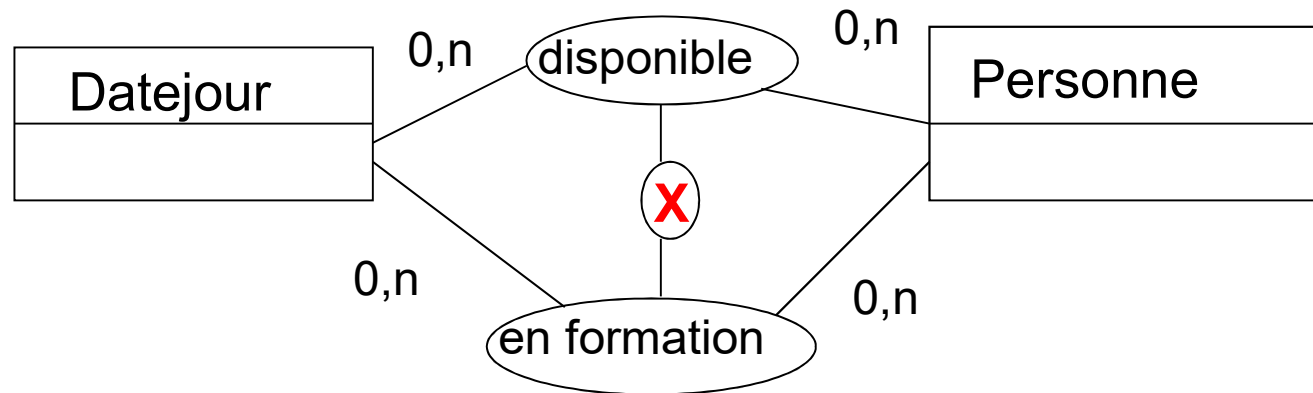
Si l'entité E participe à l'association A1, elle ne peut participer à l'association A2.



Un Article est soit acheté auprès d'un fournisseur, soit figure dans le Stock

## 2.c Exclusion de participation entre associations

Il y a exclusion de participation entre associations si la participation des entités à l'association A1 exclut leur participation à l'association A2.



Une personne à une même date ne peut pas figurer simultanément dans les deux associations: *disponible* et *en formation*.

MCD »»»»» SGBD

**DB2**

InterBase

**Informix**

Ingres

**MS Access**

MySQL

**MS SQL Server**

**Oracle**

PostgreSQL

Sybase ASE, ASA