

Petit guide d'analyse des données à l'aide de la méthode MERISE

L'analyse des données constitue le point de passage obligé de toute conception d'application mettant en œuvre un SGBDR (système de gestion de base de données relationnelle).

La méthode MERISE, basée sur le modèle entité-association, est un outil simple et efficace, très répandu chez les développeurs français. La plupart des bases de données micro pour PC (dBase, Paradox, Foxpro, Access...) sont imprégnées de cette technique pour montrer les relations entre les tables au sein d'une base.

Dans ce cours, nous allons découvrir les bases et principes fondamentaux de MERISE à l'aide d'exemples et de cas concrets.

Plan

①	Petite histoire de la méthode MERISE	les travaux de Peter CHEN et d'Hubert TARDIEU
②	Élément de base du modèle entité-association	du vocabulaire et comment on le représente graphiquement
A	Les entités	des ensembles
B	Les attributs	des caractéristiques
C	Les associations	ou relations
D	Les cardinalités	ou "combien" ?
E	Clef d'une entité	la base de la relation
	Discussion sur la qualité d'une clef	ou l'on voit que n° de sécu et plaque minéralogique sont de fausses bonnes idées
	La technique de la double clef	ou la voie la plus raisonnable
③	Topologie des associations	la forme des liens
④	Attributs d'associations	pour aller plus loin
⑤	Passage du schéma entité-association (MCD) à la construction des tables : Modèle Physique des Données (MPD)	de la théorie à la pratique
A	Transformation des entités	passer de l'entité à la table
A.1	Relation de type 1:1	la voie de la simplicité
A.2	Relation de type 1:n	maître et esclave
A.3	Relation de type n:m	plusieurs à plusieurs
B	Ou placer les attributs d'association ?	là ou c'est le mieux !
⑥	Conseils divers	... et généreux pour aller plus loin
A	Généralisation (héritage)	ramifier les espèces
B	Personnalisation	ou les sous modèles
C	Regroupement d'entités	un truc à connaître pour éviter la redondance
⑦	Exemples de MCD	pour mieux comprendre
A	Agence de location de films vidéo	ou encore pour les prêter à des copains
B	Location d'appartements pour une agence immobilière	sérieux, s'abstenir !
	BIBLIOGRAPHIE	pour poursuivre l'étude MERISE

① Petite histoire de la méthode MERISE

Le modèle entité association est un outil (et une technique) d'analyse permettant de construire des schémas théoriques de raisonnement sur des applications tournant avec les bases de données relationnelles. Pour mettre en œuvre cette technique, en connaître les possibilités et les limites, il n'est pas nécessaire de posséder une base de donnée.

Bien connaître les règles simples des schémas entités-associations (aussi appelé entités-relations) permet d'affiner petit à petit une application apparemment simple, sans avoir besoin de la programmer, et par conséquent d'économiser du temps de conception tout en obtenant une plus grande souplesse au niveau de l'analyse.

Il existe des logiciels permettant de construire des schémas entités-associations et d'en analyser les conséquences logiques, puis de construire les tables associées aux modèles de manière entièrement automatique. Ces logiciels sont appelés AGL (atelier de génie logiciel) ou CASE suivant leur puissance. Les logiciels TRAMIS, AMC*Designor, SELECT... en sont des exemples.

Le modèle entité-association est apparu dans les travaux des chercheurs, entre 1972 et 1975 lors des travaux du français MOULIN puis de TARDIEU, TEBOUL... etc. Il a été rendu célèbre dans le monde entier par l'américain Peter CHEN, à la suite d'une publication intitulée "The Entity-Relationship Model" (ACM, Transaction on Database Systems, 1976).

A ce jour tous les spécialistes français et/ou latins du domaine de l'analyse orientée base de données se servent de ce modèle comme outil de communication des applications SGBDR. Il est présent de manière transparente ou plus visible, dans la plupart des logiciels de construction d'applications de bases de données comme ACCESS, PARADOX, ORACLE, SQL Server, Informix, Ingres, Sybase...

Il n'est en revanche pas adapté aux bases de données purement objet comme O2 de Ardent Software... même si l'on admet la nouvelle dérive de MERISE orientée objet !

② Élément de base du modèle entité association

Le modèle entité association est constitué de deux éléments de base :

- Les entités, qui sont des regroupements d'informations, et possèdent des attributs (caractéristiques)
- Les associations qui sont les liens logiques entre les entités (et sont quantifiées par des cardinalités)

A Les entités

Ce sont des regroupements d'informations. Les informations contenues dans les entités (informations que l'on appelle "attributs") doivent être des informations variables, mais communes à une même classe d'objets.

Par exemple, si l'on considère l'entité "être humain" les informations communes aux être humains peuvent être :

- le nom,
- le ou les prénoms,
- la date de naissance,
- le lieu de naissance,
- le sexe,
- l'adresse, etc.

On considère souvent qu'il s'agit de "classes" d'entités.

Une entité donnée peut elle-même être constituée de sous-classes.

Par exemple, un être humain donné peut habiter au même endroit qu'un autre (si deux personnes vivent sous le même toit parce qu'ils sont mariés). Dans ce cas, l'adresse constitue une sous-classe de l'entité "être humain", c'est à dire une nouvelle entité à part entière.

D'un autre côté, il arrive souvent que plusieurs personnes résident au même endroit, sans même se connaître (cas d'un immeuble collectif par exemple). Dans ce cas on peut considérer l'adresse, comme une entité et la décrire de la manière suivante :

- Pays
- Région
- Département
- Rue etc.

On schématise une entité par un rectangle.

Exemple :



B Les attributs

Les attributs sont les caractéristiques décrivant les entités et doivent être représentés comme une liste de mots, la plus simple possible, dans le cadre de l'entité correspondante. On devra préciser le type des données attendues pour chaque attribut.

Exemple :

ETRE HUMAIN		ADRESSE	
Nom	A32	Ligne adresse 1	A32
Prénom	A25	Ligne adresse 2	A32
Date de naissance	D	Ligne adresse 3	A32
Lieu de naissance	A32	Code postal	A7
Sexe	BL	Ville	A32
		Pays	A32

Les types associés aux attributs sont les suivants :

D	Date
Annn	Caractères de longueur nnn
BL	Booléen (vrai / faux)
T	Temps
DT	Date Temps
N	Nombre
S	(Smallint) entier court
I	(Integer) entier

Etc.

C Les associations

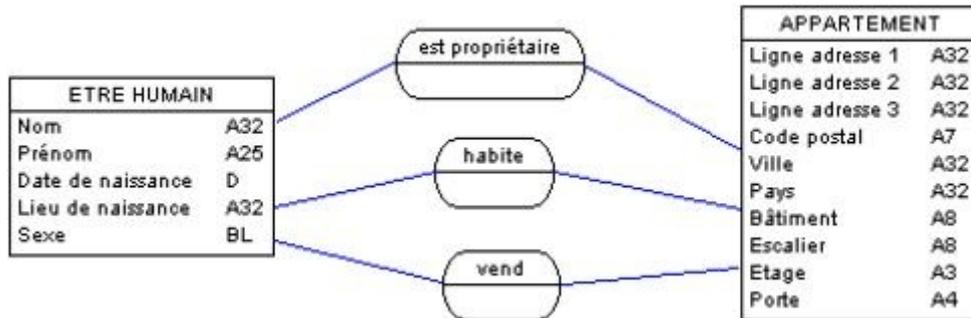
Ce sont des liaisons logiques entre les entités.

Elles peuvent être de nature factuelle, ou de nature dynamique.

Par exemple, une personne peut acheter un objet (action d'acheter), mais si l'on considère qu'une personne est propriétaire d'un objet, alors l'association entre l'objet et cette personne est purement factuelle.

Les associations se représentent dans une ellipse (ou un rectangle aux extrémités rondes), reliée par des traits aux entités qu'elles lient logiquement.

Exemple :



D Les cardinalités

Les cardinalités, au sens arithmétique du terme, permettent de dénombrer les éléments de l'entité d'arrivée en relation avec un élément de l'entité de départ, et vice versa.

Exemple : considérons le cas de l'association "habite" et les deux entités "être humain" et "appartement" du schéma précédent :

les cardinalités minimales et maximales sont les suivantes :

sens "être humain" vers "appartement" : 1 (minimum) et 1 (maximum)

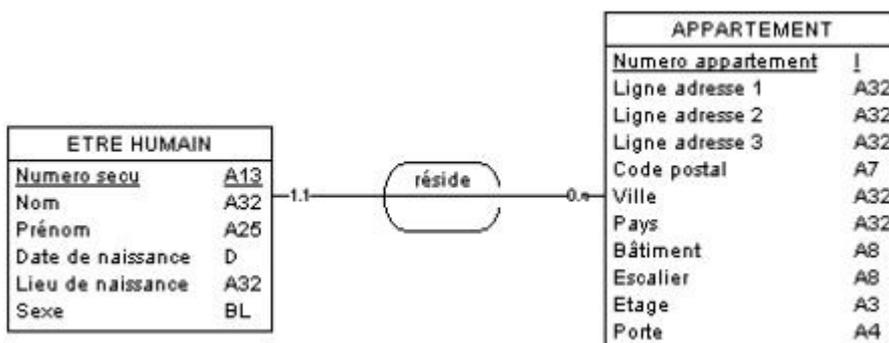
sens "appartement" vers "être humain" : 0 (minimum) et n (maximum)

Ce qui signifie que dans cette modélisation un être humain réside dans un appartement et un seul à la fois, mais qu'un appartement peut se trouver vide ou être pourvu de plusieurs résidents.

Bien entendu tout être humain ne réside pas forcément dans un appartement, ce peut être dans une maison, à l'hôtel ou à la belle étoile (SDF...). Mais il convient de se concentrer sur ce que l'on doit modéliser et non sur l'univers entier.

En outre nous avons convenu qu'un même être humain ne pouvait résider dans plusieurs appartements à la fois (notion de « résidence principale » par exemple).

On note les cardinalités de chaque côté de l'association, sur les traits faisant la liaison entre l'association et l'entité.



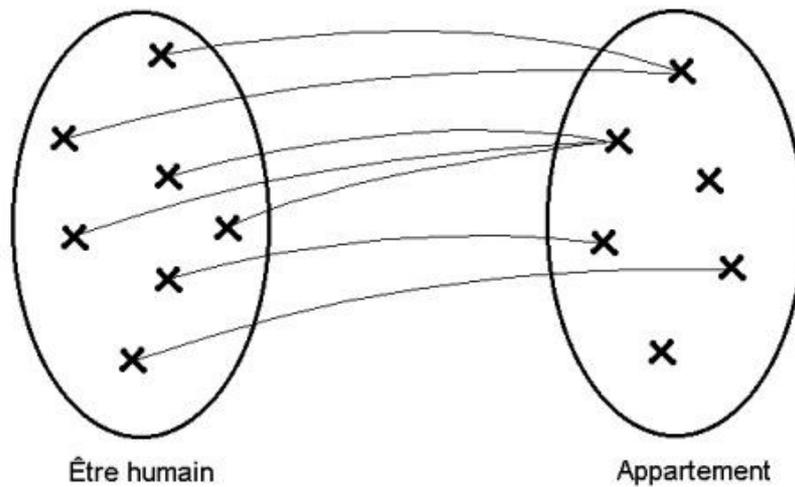
Petit guide de la méthode MERISE

Dès cet instant, on peut en déduire le type de relation parmi les types suivants : : 1/1, 1/n ou n/1 ou encore n/m. Ici c'est une relation de type 1,n.

Pour déduire le type d'une relation, il suffit de récupérer les cardinalités maximales des deux côtés de l'association, sans tenir compte de leur valeur exacte.

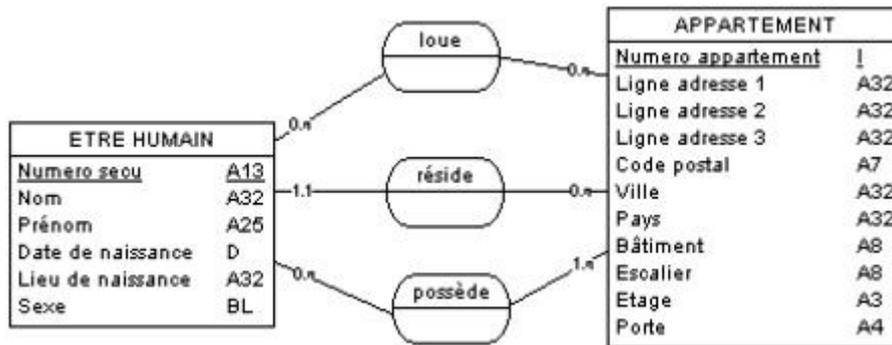
On peut s'y aider à l'aide des diagrammes de Wen et des notions de mathématiques ensemblistes, bien connus de ceux qui ont pratiqués les maths modernes dès leur plus jeune âge...

Entre alors en jeu les notions de bijection, interjection, surjection...



Attention ! Des relations différentes entre mêmes entités peuvent posséder des cardinalités différentes; c'est même souvent le cas.

Exemple :



La relation « loue » est de type n : m

La relation « réside » est de type 1 : n

La relation « possède » est de type n : m

E Clef d'une entité

C'est un attribut (ou un ensemble d'attributs) qui permet (tent) de distinguer un élément de l'entité de manière unique et sans aucune ambiguïté par rapport à l'ensemble des autres éléments, et à l'univers de tous les éléments qui peuvent entrer un jour ou l'autre dans cette entité.

Par exemple la clef de l'entité "être humain" pourrait être le nom. Mais comme le cas d'homonymie est assez fréquent, surtout lorsque l'on manipule des fichiers volumineux, alors cet attribut constitue une mauvaise clef en général.

En revanche, il n'est pas impossible que la clef d'une entité soit composée de plusieurs attributs. Par exemple, la clef de l'entité "être humain" pourrait être le nom et le prénom. Cependant il n'est toujours pas impossible d'avoir deux personnes dont le nom et le prénom soient identiques...

On note qu'un attribut est une clef en le soulignant dans le schéma entité association. Si c'est une clef composée, alors plusieurs entités seront soulignées.

Exemple :

ETRE HUMAIN	
<u>Nom</u>	A32
<u>Prénom</u>	A25
<u>Date de naissance</u>	D
Lieu de naissance	A32
Sexe	BL

APPARTEMENT	
<u>Ligne adresse 1</u>	A32
<u>Ligne adresse 2</u>	A32
<u>Ligne adresse 3</u>	A32
<u>Code postal</u>	A7
<u>Ville</u>	A32
<u>Pays</u>	A32
<u>Bâtiment</u>	A8
<u>Escalier</u>	A8
<u>Etage</u>	A3
<u>Porte</u>	A4

On voit ici que dans le cas de l'entité "appartement" tous les attributs sont utilisés pour composer la clef. Cette clef naturelle n'étant pas pratique, il est plus judicieux de créer un nouvel attribut qui servira expressément de clef à l'association.

Pour l'entité "être humain", on pourrait se servir du numéro de sécurité sociale (plus exactement du numéro INSEE) , comme clef de l'entité. En revanche, pour ce qui est de l'entité "appartement" il est conseillé de créer un nouvel attribut clef qui serait, par exemple, un numéro.

Exemple :

ETRE HUMAIN	
<u>Numero secu</u>	A13
<u>Nom</u>	A32
<u>Prénom</u>	A25
<u>Date de naissance</u>	D
<u>Lieu de naissance</u>	A32
<u>Sexe</u>	BL

APPARTEMENT	
<u>Numero appartement</u>	I
Ligne adresse 1	A32
Ligne adresse 2	A32
Ligne adresse 3	A32
Code postal	A7
Ville	A32
Pays	A32
Bâtiment	A8
Escalier	A8
Etage	A3
Porte	A4

Pour une entité de type « Voiture » il pourrait être fait usage de l'immatriculation du véhicule comme clef de l'entité.

MAIS... lisez attentivement ce qui suit !

DISCUSSION SUR LA QUALITÉ D'UNE CLEF

De manière générale il convient de limiter les clefs composées. Chaque fois que l'on aura le choix entre la création d'une clef numérique, et une clef naturelle mais composée, il sera préférable de créer une clef numérique, à de très rares exceptions près.

En effet les SGBDR sont plus « à l'aise » lorsqu'ils ont à manipuler des clefs purement numériques.

De plus une clef est un concept purement informatique.

Par exemple le n° de sécurité sociale est une mauvaise clef en générale : un individu étranger qui arrive sur le sol français est doté d'une immatriculation provisoire et ne connaît son numéro définitif que plusieurs mois après avoir rempli les conditions requises par l'administration.

Par exemple l'immatriculation d'un véhicule est une mauvaise clef : en effet, du fait de la fiscalité sur les véhicules à moteur (et en particulier les vignettes), les sociétés n'hésitent pas à faire immatriculer leur parc de véhicules dans le département où les taxes sont les moins élevées (le 51). Cette immatriculation peut donc être amenée à changer.

Or toute clef évolutive est un danger pour le système informatique : si la valeur de la clef change, nous verrons qu'il faut la modifier dans tous les fichiers dans laquelle elle est référencée. Il est vrai que certains SGBDR autorise automatiquement ce genre de manœuvre, mais cette automatisation très contraignante pénalise lourdement le fonctionnement du SGBDR.

On veillera donc à prendre une clef totalement indépendante des attributs ordinaires de l'entité.

De plus une clef a intérêt à être la plus courte possible afin que son stockage soit limité à quelques octets dans les fichiers et donc le temps de recherche le plus rapide possible.

Le plus simple consiste donc à introduire dans le descriptif de l'entité une clef strictement « informatique » qui se résumera en général à un numéro (entier long) que l'on pourra incrémenter automatiquement.

Parfois, il arrive qu'une clef ait été créée spécifiquement pour les besoins des systèmes informatiques. C'est en particulier le cas des indicatifs téléphoniques (33 pour la France) ou des code de pays (F pour France). Dans ce dernier cas on devra alors utiliser la clef commune, comme clef de l'entité.

TECHNIQUE DE LA DOUBLE CLEF

Une technique éprouvée consiste à introduire une double clef dans toutes les tables : la clef « informatique » et une clef « utilisateur ».

La clef informatique est l'index primaire de la table et doit posséder les caractéristiques suivantes :

purement numérique (par exemple un entier long)

unique bien entendu

obligatoire

sans mise à jour en cascade

avec intégrité référentielle

générée automatiquement

invisible pour l'utilisateur.

La clef utilisateur doit être assez « souple », c'est à dire posséder dans la mesure du possible, les caractéristiques suivantes :

Index secondaire unique

Obligatoire

Utiliser un jeu de caractère réduit s'il s'agit d'un format alpha (par exemple les 26 lettres majuscules de l'alphabet et les chiffres de 0 à 9)

Limité à une faible taille (16 octets - 16 caractères maximum)

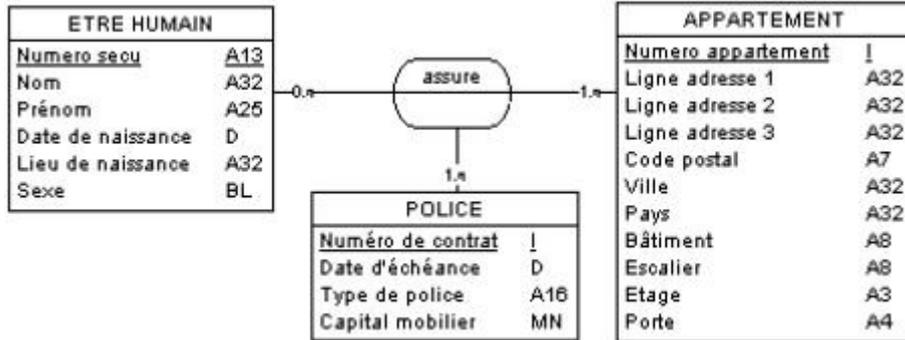
La présentation des données devant toujours s'effectuer suivant l'ordre établis par la clef utilisateur à défaut d'autre spécification

③ Topologie des associations

La plupart des associations sont de nature binaire, c'est à dire composées de deux entités mise en relation par une ou plusieurs associations.

C'est le cas par exemple de l'association "est propriétaire" mettant en relation "être humain" et "appartement". Cependant il arrive qu'une association concerne plus de deux entités (on dit alors qu'il s'agit d'association "n-aires").

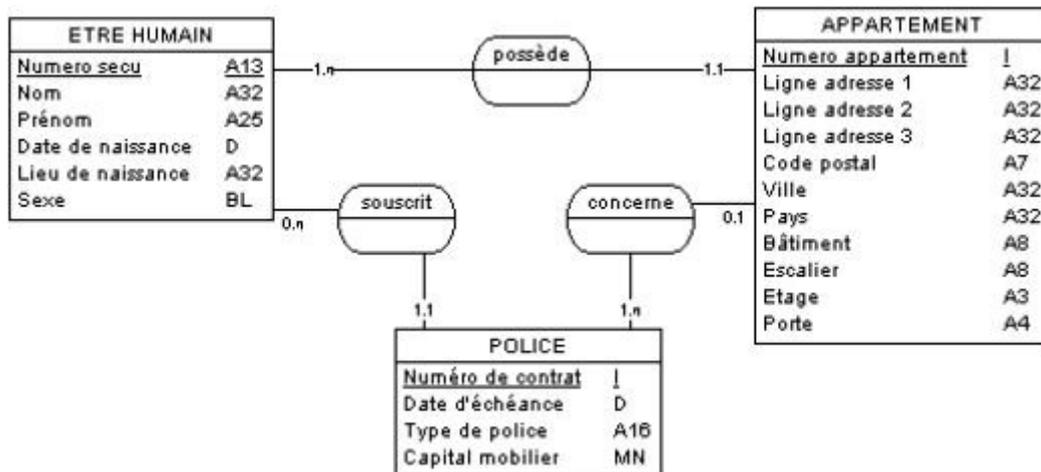
Exemple :



Mais dans ce cas il y a de grandes difficultés pour exprimer les cardinalités.

On aura tout intérêt à essayer de transformer le schéma de manière à n'obtenir que des associations binaires.

Exemple :



④ Attributs d'association

Il arrive parfois que l'on soit obligé de munir d'attributs des associations.

Exemple : considérons par exemple, que nous voulons modéliser les relations existant entre les entités "client", "commande" et "article" :

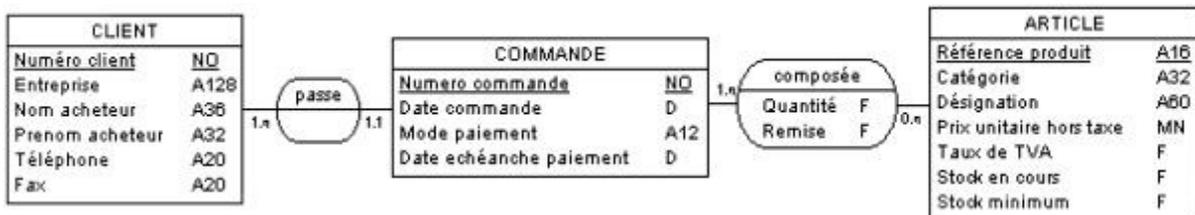


Mais comment dans ce schéma introduire l'attribut "quantité" et plus encore l'attribut "réduction" dont on voudrait qu'il puisse s'appliquer à chacun des articles d'une commande de manière différente ?

En effet si l'on introduit l'attribut quantité à l'entité COMMANDE, chaque ligne de la commande se vera dotée de la même quantité...

D'autre part si l'on introduit l'attribut quantité à l'entité ARTICLE alors chacun des article se vera doté de la même quantité quelque soit la commande...

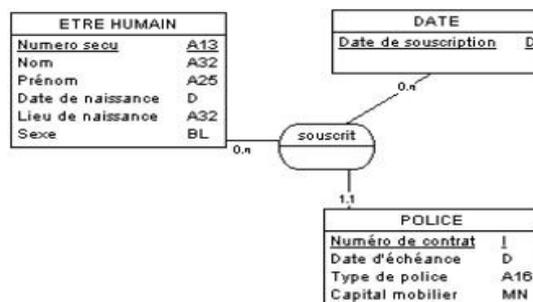
La solution est de pourvoir l'association "composée" des attributs "quantité" et "réduction" :



Il arrive dans certains cas que l'attribut "date" soit d'une importance capitale, notamment dans les applications SGBDR portant sur la signature de contrats à échéance ou dans la durée (assurance par exemple).

Il n'est pas rare alors que le seul attribut "date" constitue à lui seul une entité.

Exemple :



On appelle alors cela une entité temporelle. Une entité temporelle possède souvent un seul attribut, mais dans le cas où elle possède plusieurs attributs (année, mois, jour, heure, minute, seconde...), l'ensemble de ces attributs constitue alors la clef de l'entité.

Mais dans ce cas on peut aussi retirer cette entité et introduire la date en tant qu'attribut de l'association "souscrit".

⑤ Passage du schéma entité-association (MCD) à la construction des tables : Modèle Physique des Données (MPD)

Ce que nous venons de voir concerne l'analyse conceptuelle des données, c'est à dire un niveau d'analyse qui s'affranchi de toutes les contraintes de la base de données sur lequel va reposer l'application. Une fois décrit sous forme graphique, ce modèle est couramment appelé MCD pour "Modèle Conceptuel des Données".

Dès lors, tout MCD peut être transformé en un MPD ("Modèle Physique des Données") c'est à dire un modèle directement exploitable par la base de données que vous voulez utiliser...

Tout l'intérêt de cet outil d'analyse est de permettre de modéliser plus aisément les relations existant entre les entités et d'automatiser le passage du schéma muni d'attributs aux tables de la base de données pourvues de leurs champs.

Voici maintenant les règles de base nécessaire à une bonne automatisation du passage du MCD au MPD :

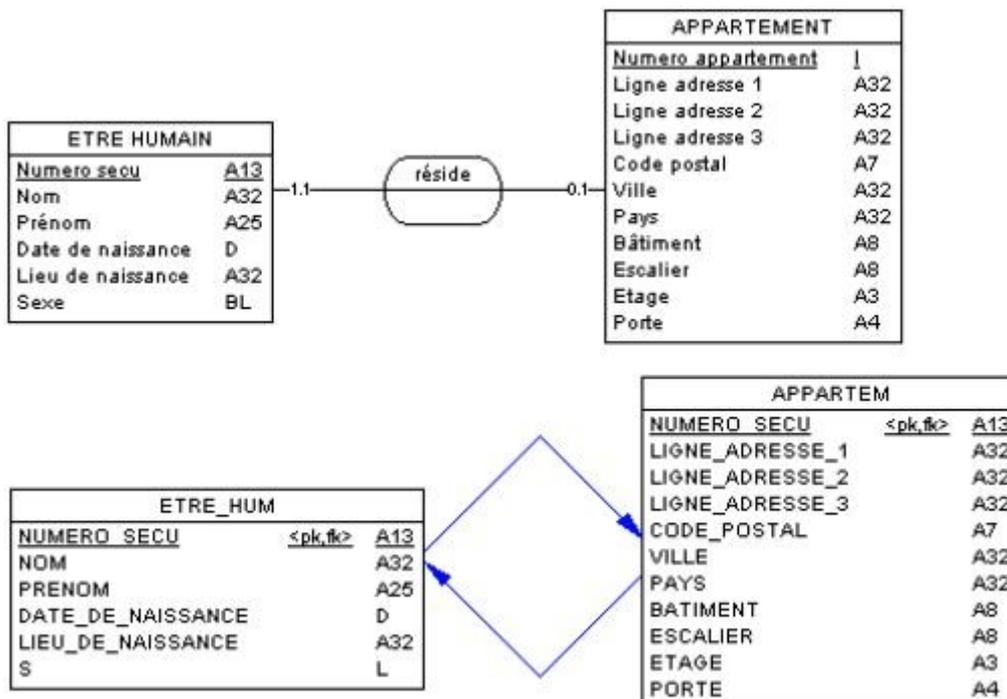
A Transformation des entités

Règle n°1 : toute entité doit être représentée par une table.

A.1 Relations de type 1:1

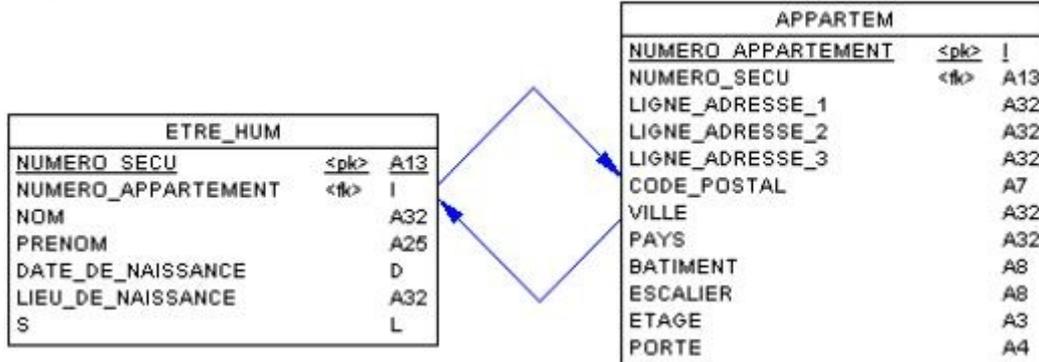
Règle n°2 : Dans le cas d'entités reliées par des associations de type 1:1, les tables doivent avoir la même clef.

Exemple :



NOTA : on aurait pu choisir l'autre clef comme clef commune, à savoir le n° d'appartement.

Bien que certains systèmes proposent une autre solution :

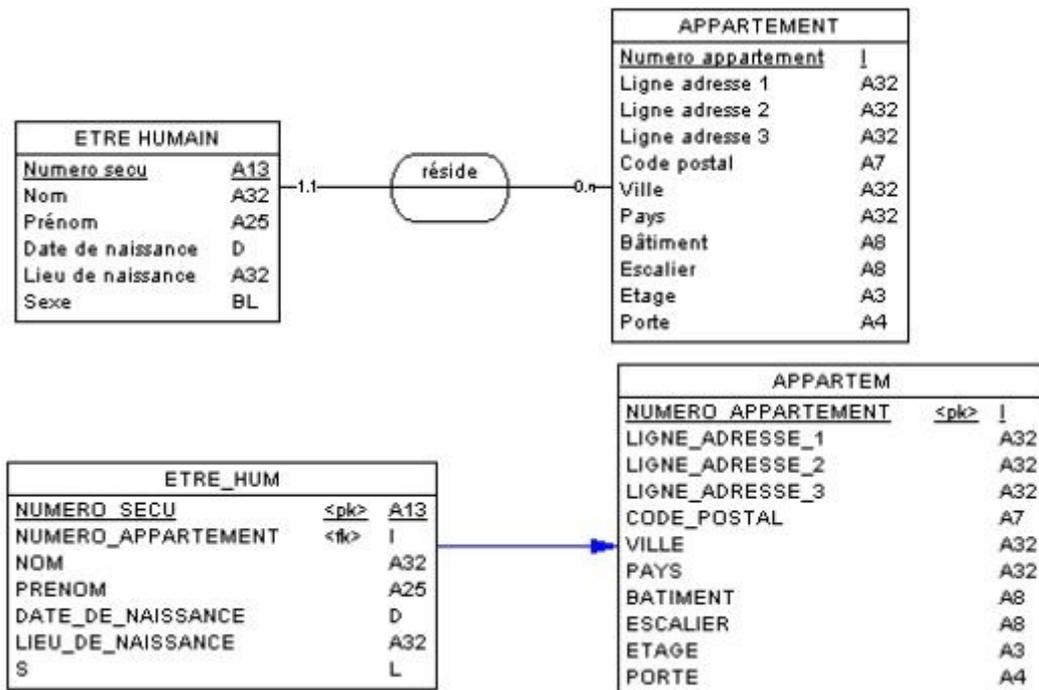


Dans ce cas une étude approfondie de la solution à adopter est nécessaire, mais ce type de relation est en général assez rare et peu performante...

A.2 Relations de type 1:n

Règle n°3 : Dans le cas d'entités reliées par des associations de type 1:n, chaque table possède sa propre clef, mais la clef de l'entité côté 0,n (ou 1,n) migre vers la table côté 0,1 (ou 1,1) et devient une clef étrangère (index secondaire).

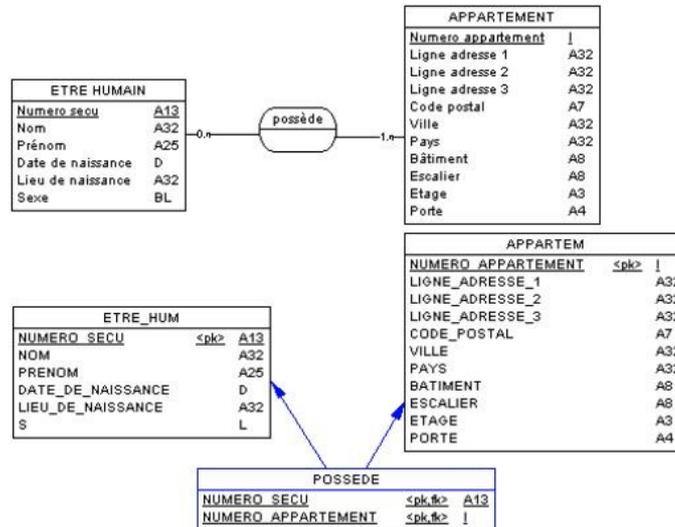
Exemple :



A.3 Relations de type n:m

Règle n°4 : Dans le cas d'entités reliées par des associations de type n:m, une table intermédiaire dite table de jointure, doit être créée, et doit posséder comme clef primaire une conjonction des clefs primaires des deux tables pour lesquelles elle sert de jointure.

Exemple :

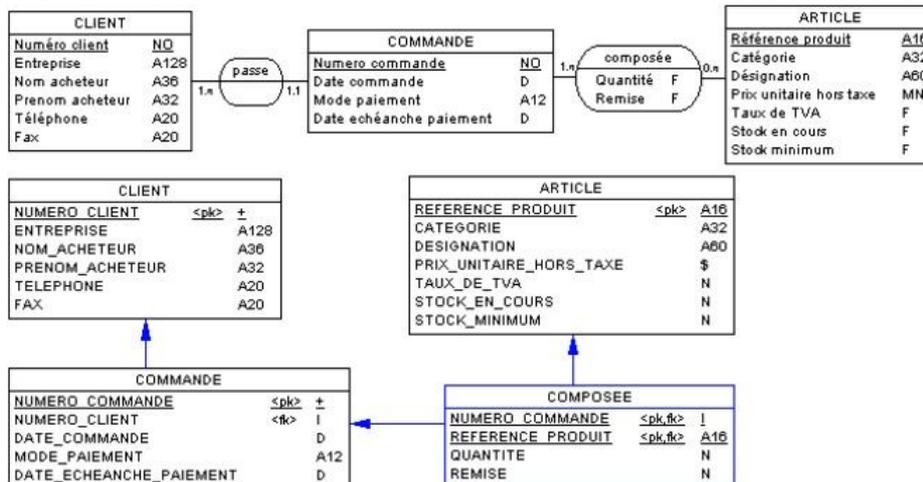


B Ou placer les attributs d'association ?

Règle n°5 : Cas des associations pourvues d'au moins un attribut :

- si le type de relation est n:m, alors les attributs de l'association deviennent des attributs de la table de jointure.
- si le type de relation est 1:n, il convient de faire glisser les attributs vers l'entité pourvue des cardinalités 1:1.
- si le type de relation est 1:1, il convient de faire glisser les attributs vers l'une ou l'autre des entités.

Exemple : Pour synthétiser toutes ces règles, voici un exemple de modélisation d'une application. En l'occurrence il s'agit d'un service commercial désirant modéliser les commandes de ses clients.

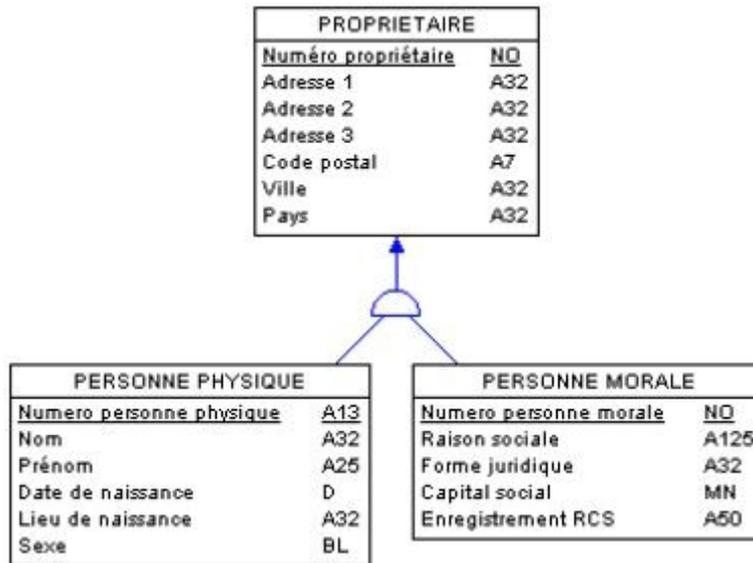


⑥ Conseils divers

A Généralisation (héritage)

Dans le schéma ci-dessous, les entités "Personne physique" (des êtres humains) et "Personne morales" (des sociétés, associations, collectivités, organisations...) sont généralisées dans l'entité "Propriétaires". On dit aussi que l'entité "Propriétaire" est une entité parente et que les entités "Personne morale" et "Personne physique" sont des entités enfants, car il y a une notion d'héritage...

Exemple :

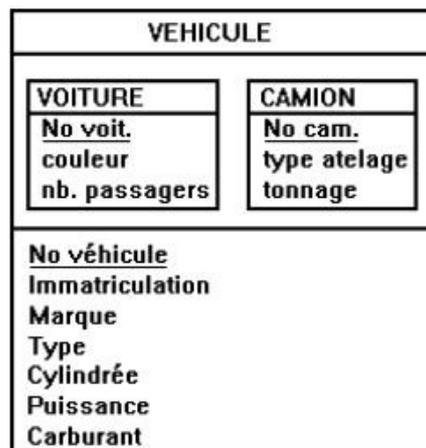


Par exemple une entité "Etre humain" est une généralisation pour toute entité faisant appel à une personne, comme les entités "Etudiant", "Client", "Artiste", "Souscripteur", "Patient", "Assujetti"...

On les appelle aussi "entités-génériques".

Certains atelier de modélisation représentant les données sous la forme d'entités « encapsulés ».

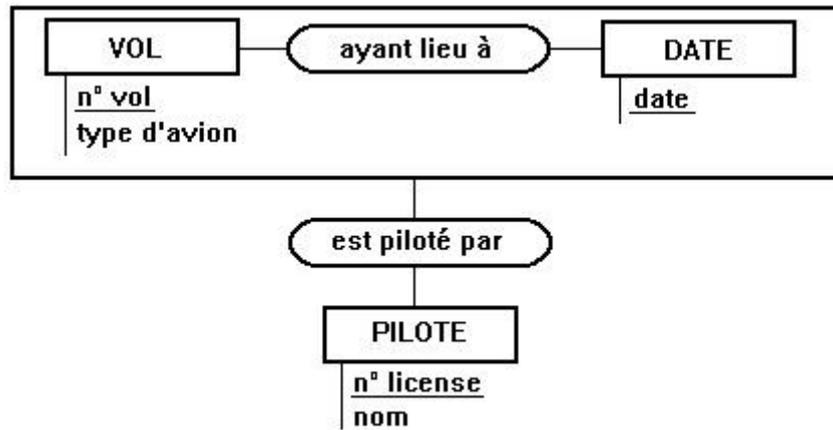
Exemple :



B Personnalisation

Une personnalisation est un regroupement dans une super entité de plusieurs entités munies d'une ou de plusieurs associations.

Par exemple, une compagnie d'aviation proposant des vols peut modéliser le planning des pilotes par le schéma suivant :



C Regroupement d'entités

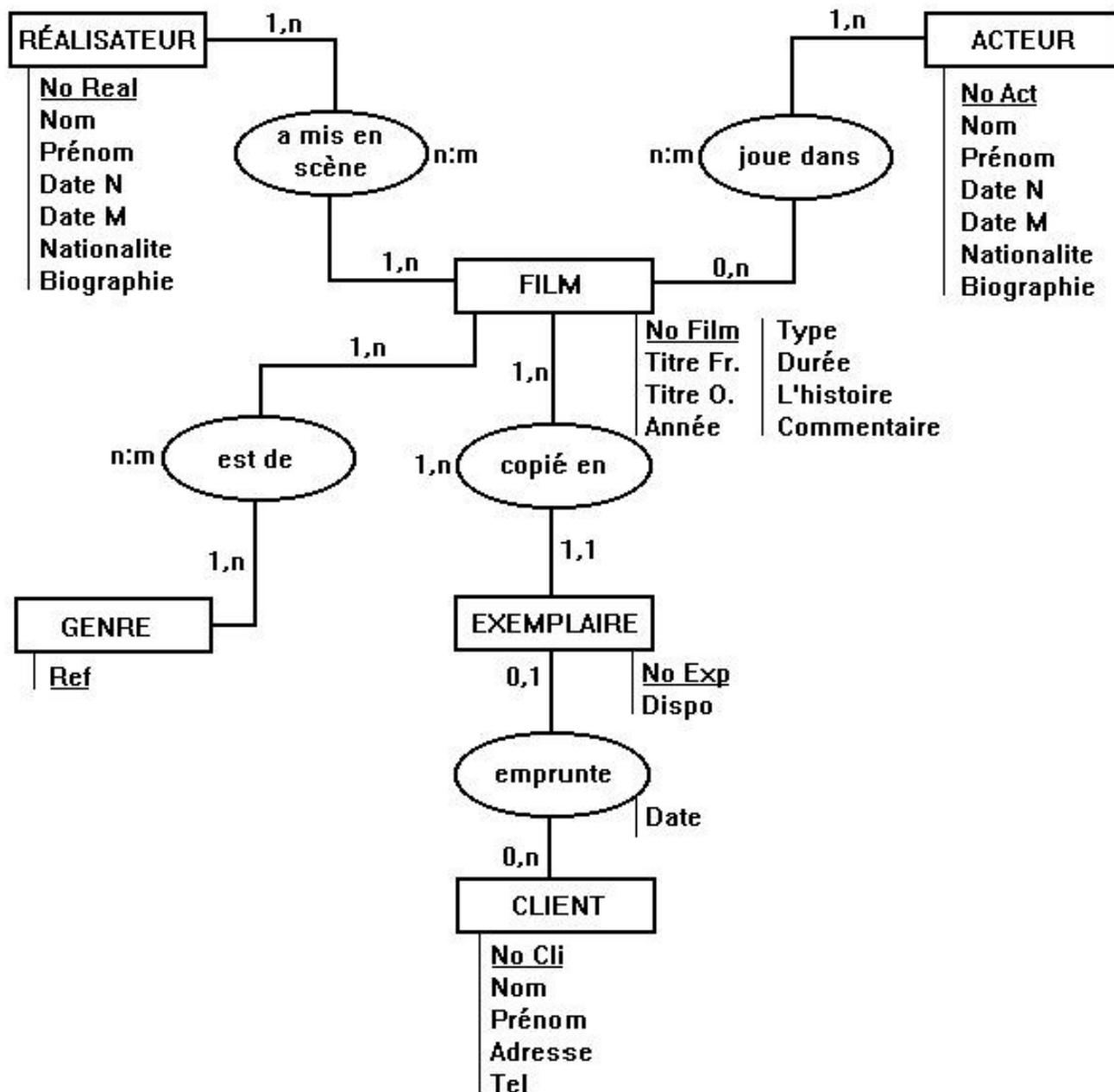
Comme toute technique, le schéma entité-association possède des limites et des contraintes que seuls l'expérience et le bon sens peuvent permettre d'éliminer.

Il arrive parfois que certaines entités apparaissent comme redondantes. Dans ce cas, et pour gagner de la place en matière de stockage de l'information, il convient de regrouper ces entités dans une seule et même table du SGBDR en ajoutant un champ supplémentaire à cette table de manière à permettre de distinguer les entités du schéma théorique.

Par exemple si l'on désire modéliser une gestion de compact-disc on peut créer une entité "Compositeur" et une entité "Interprète". Mais on constate qu'une grande majorité de compositeurs sont leurs propres interprètes, ce qui signifie qu'une même personne peut se trouver présente dans les deux entités. Pour résoudre ce problème il suffit de construire une seule table pour les deux entités (par exemple une table "MUSICIEN") et d'y ajouter un champ permettant de distinguer le type de "musicien" : compositeur ou interprète ou les deux.

⑦ Exemples de MCD

A Agence de location de films vidéo

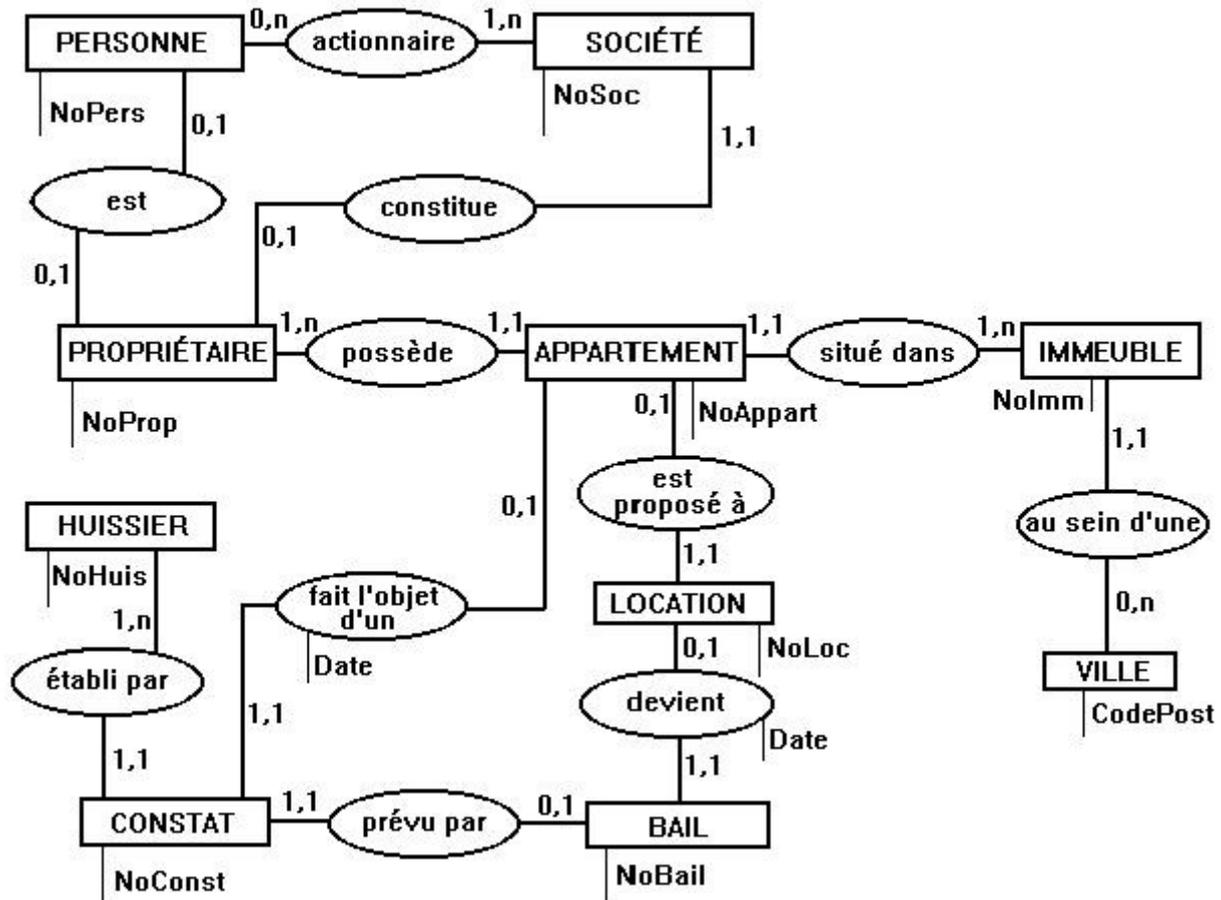


Lors de la réalisation de la base de données, les entités RÉALISATEUR et ACTEUR peuvent être regroupées en une seule table car il y a un nombre non négligeable de réalisateurs qui sont acteurs, et vice-versa. Dans ce cas, un champ d'un seul caractère permettra de faire la différence entre un réalisateur pur, un acteur pur et un acteur réalisateur.

Notez aussi les cardinalités entre les entités ACTEUR et FILM, en effet, un film d'animation ne possède aucun acteur.

Dans l'entité EXEMPLAIRE figure un attribut "dispo" permettant de savoir si l'exemplaire n°X d'un film est disponible ou en cours d'emprunt.

B Location d'appartements pour une agence immobilière



NOTA : pour simplifier l'écriture, ne figurent dans ce schéma que les attributs clefs ou les attributs d'associations.

BIBLIOGRAPHIE

- Bases de Données et Modèles de Calcul - Outils et Méthodes pour l'Utilisateur - Jean-Luc HAINAUT - InterEditions 1994
- Conception de bases de données : du schéma conceptuel au schéma physique - GALACSI - Dunod 1989
- La méthode MERISE, principes et outils (Tome 1 & 2) - TARDIEU, ROCHFELD, COLLETTI - Les éditions d'organisation 1986
- Modélisation dans la conception des systèmes d'information - ACSIOME - Masson 1990
- Apprendre et pratiquer MERISE - J. GABAY - Masson 1993
- MERISE, vers une modélisation orientée objet – José MOREJON – Les Editions d'Organisation 1994
- Maîtriser les bases de données – Georges GARDARIN – Eyrolles 1993
- Introduction aux bases de données, 6e édition– Chris J. DATE – Thomson International Publishing 1998
- Concepts fondamentaux de l'informatique – Alfred AHO, Jeffrey ULLMAN – Dunod 1993
- AMC*Designor, mise en œuvre de MERISE – Gilles GUEDJ – Eyrolles 1996

WEBOGRAPHIE

- Memo Merise (toute la théorie de MERISE, incluant les modèles de traitement de données et les formes normales) :
 - <http://perso.wanadoo.fr/matthieu.vidal/default.htm>
- Merise par Stéphane LAMBERT :
 - <http://www.sam-mag.com/archives/merise.htm> (1)
 - <http://www.sam-mag.com/archives/mldr.htm> (2)
 - <http://www.sam-mag.com/archives/creerbase.htm> (3)
 - <http://www.sam-mag.com/archives/modeliser.htm> (4)