

Table of Contents

I. NOTIONS IMPORTANTES.....	3
II. MODELE RELATIONNEL	4
1) Introduction	4
A) 3 composantes	4
a) Les concepts relationnels.....	4
b) Algèbre relationnel	4
c) Contraintes d'intégrité relationnelle.....	4
B) Rappels mathématiques.....	4
a) Produit cartésien.....	4
b) Relation binaire.....	4
2) Concepts Relationnels.....	5
A) Domaine sémantique.....	5
B) Relations / Tuples.....	5
C) Attributs.....	6
D) Clé primaire.....	7
E) Domaine primaire.....	7
F) Clé étrangère	7
G) Schéma relationnel.....	7
H) Autres concepts	8
a) Le degré et la cardinalité d'une relation.....	8
b) Relations dynamiques et statiques.....	8
c) Attributs comparables ou compatibles.....	8
3) Contraintes d'intégrités relationnelles.....	9
A) Contraintes d'intégrité statique.....	9
a) Contrainte de domaine.....	9
b) Contrainte de relation.....	9
c) Contrainte de référence.....	9
B) Contraintes dynamiques.....	9
4) Algèbre relationnel	10
A) Généralités.....	10
a) Opérateurs relationnels.....	10
Relation unaire.....	10
Relation binaire.....	10
b) Opérateurs ensemblistes.....	10
B) Opérateurs relationnels.....	10
a) Projection.....	10
b) Sélection.....	11
c) Jointure.....	12
L'équijointure.....	13
Jointure naturelle.....	14
Jointure par \ltimes	14
Autojointure.....	14
d) La division.....	14

C)Opérateurs ensemblistes.....	15
III.DEPENDANCES FONCTIONNELLES.....	17
1)Notations	17
2)Définitions.....	17
3)Axiomes d'Amstrong	18
a)La réflexivité.....	18
b)L'augmentation.....	19
c)La transitivité.....	19
d)La pseudo-transitivité.....	19
e)L'union.....	19
f)La décomposition.....	19
4)Autres concepts	19
a)Dépendance fonctionnelle minimale.....	19
b)Clé candidate.....	20
c)Clé candidate minimale	20
d)Clé primaire.....	21
e)Attributs non-clé primaire / non clé candidate minimale.....	21
IV.NORMALISATION DES RELATIONS.....	22
1)Objectif.....	22
2)Formes normales.....	23
A)1NF.....	23
B)2NF.....	24
C)3NF.....	25

I. NOTIONS IMPORTANTES

Définition :

Une base de données est un réservoir structuré de données partageables par plusieurs utilisateurs

Une base de données doit :

- Être organisée , afin d'optimiser le stockage
- Être consultable
- Être mise à jour sans perte de données :
 - Ajouter
 - Modifier
 - Supprimer

Modèle de données :

- Concepts pour définir les données
- Des opérateurs de manipulation des données
- Des contraintes d'intégrités : des règles de cohérence

Différents modèles :

- 1960 : IBM : Modèle hiérarchique : n'a pas fonctionné
- 1960 : IBM : Modèle réseau : n'a pas fonctionné
- 1970 : IBM et Codd : modèle relationnel basé sur la théorie des ensembles , toujours utilisé depuis

SGBDR : Système de Gestion de Base de Données Relationnel

SGBD : implantation du modèle de données , environnement de développement d'application BD

Concurrents : Oracle , MySQL , Access , Systèmes Relationnels , PostGre SQL , sql lite

II. MODELE RELATIONNEL

1) Introduction

A) 3 composantes

a) *Les concepts relationnels*

Domaine , Relations , Tuples , Attributs , Clé primaire , Clé étrangère , relation statique , relation dynamique , schéma relationnel .

b) *Algèbre relationnel*

Projection , sélection , jointure , division , union , intersection ,différence .

c) *Contraintes d'intégrité relationnelle*

Intégrité de domaine , inégibilité de relation , intégrité de références

B) Rappels mathématiques

a) *Produit cartésien*

Soient A et B deux ensembles tels que :

$$A \times B = \langle (a, b) | a \in A, b \in B \rangle$$

Exemple :

$$A = \langle 1, 2, 3 \rangle \quad B = \langle 4, 5 \rangle$$

$$A \times B = \langle (1, 4), (1, 5), (2, 4), (2, 5), (3, 4), (3, 5) \rangle$$

b) *Relation binaire*

Soient A et B , deux ensembles $R \subseteq A \times B$ est une relation entre les éléments de A et de B

R contient les couples de $A \times B$

exemple :

$$R1 = \langle (a1, b1) \rangle \quad R2 = \langle (a1, b1), (a1, b2) \rangle$$

2) Concepts Relationnels

A) Domaine sémantique

Un domaine sémantique est un ensemble de valeurs atomiques , ou simples .

Chaque domaine est caractérisé par un type de base (numérique : entier , réel ; alphanumérique : string , date ...)

Exemple :

D_NOMAV = Domaine de nom des avions

D_NOMPIL = Domaine de nom des pilotes

D_NOMAV={'A310','A330','A340','B727','B747'}

D_NOMPIL={'DUPONT','ROGER'}

Le domaine sémantique a comme objectif de vérifier la validité des comparaisons (évite de comparer D_NOMPIL avec D_NOMAV)

B) Relations / Tuples

Soient D1 , D2 , Dn , n domaines non nécessairement distincts .

Une relation R est le sous ensemble du produit cartésien de ces n domaines

$$R \subseteq D1 \times D2 \times \dots \times Dn$$

La relation en BD est une relation n-aire avec des n-uplets comme éléments

Exemple :

AVION

D_NUMAV

D_NOMAV

D_CAP

D_VILLE

$$AVION \subseteq D_NUMAV \times D_NOMAV \times D_CAP \times D_VILLE$$

Tuple : représente un avion :

(100,'A310',250,'MARSEILLE')

Représentation tabulaire d'une relation

AVION={ (100,'A310',250,'MARSEILLE'),(200,'B747',450,'PARIS') }

AVION	NUMAV	NOMAV	CAP	VILLE
	100	A310	250	MARSEILLE
	200	B747	450	PARIS

Ici ,

AVION est l'en-tête

NUMAV , NOMAV , CAP , VILLE sont les attributs

Le corps est constitué des différents tuples

Les tuples sont surlignés en jaune : les lignes du tableau

Le corps est l'extension du schéma relationnel

L'en-tête est le schéma de la relation / intention

C) Attributs

L'attribut est un rôle joué par un domaine dans une relation pour éliminer les ambiguïtés ($D_{VILLE} \times D_{VILLE}$)

$VOL \subseteq D_NUMVOL \times D_VILLE \times D_VILLE \times D_HEURE \times D_HEURE$

Attribut NUMVOL VILLE_DEP VILLE_ARR H_DEP H_ARR

Un attribut joue un rôle dans un domaine , un domaine a un type défini

Ainsi , (IT100,'PARIS','LONDRES',18,20)

Avec les attributs :

D_NUMVOL => NUMVOL

D_VILLE => VILLE_DEP

D_VILLE => VILLE_ARR

D_HEURE => HEURE_DEP

D_HEURE => HEURE_ARR

Chaque attribut est défini sur un domaine

D) Clé primaire

Une clé primaire est un attribut , ou groupe d'attributs permettant d'identifier de manière unique les tuples de la relation .

Exemple : NUMAV ne contient aucun doublon , est ainsi le domaine de clé primaire de la relation AVION

E) Domaine primaire

C'est le domaine dans lequel est défini un attribut, une clé primaire.

Exemple : D_NUMAV : domaine primaire

NUMAV : clé primaire

F) Clé étrangère

C'est un attribut ou groupe d'attribut défini sur un ou plusieurs domaines primaires

C'est un attribut , ou groupe d'attribut clé primaire dans une autre relation .

Exemple :

AVION(NUMAV,NOMAV,CAP,LOC)

VOL(NUMVOL,#NUMAV,#NUMPIL,VILLE_DEP,VILLE_ARR,HEURE_DEP,HEURE_ARR)

PILOTE(NUMPIL,NOMPIL,ADR,SALAIRE)

G) Schéma relationnel

C'est le schéma de chaque relations , la liste des attributs

La clé étrangère sert à exprimer des liens sémantiques entre les relations du schéma relationnel

C'est l'ensemble de schéma de relations liées à travers des clé primaires et étrangères .

Chaque schéma de relation est décrit par le nom de la relation et la liste de ses attributs :

AVION(NUMAV,NOMAV,CAP,LOC)

H) Autres concepts

a) *Le degré et la cardinalité d'une relation*

Le degré d'une relation est le nombre des attributs d'une relation .

Ex:

Degré(AVION)=4

La cardinalité d'une relation est le nombre de ses tuples .

Ex :

Card(AVION) =2 (voir le schéma page 5)

b) *Relations dynamiques et statiques*

-Une relation dynamique est une relation qui contient une clé étrangère .

Ex: VOL(NUMVOL,#NUMAV,#NUMPIL,VILLE_DEP,VILLE_ARR,HEURE_DEP,HEURE_ARR)

Les clé étrangères sont NUMAV et NUMPIL

-Une relation statique est une relation indépendante , qui ne contient pas de clé étrangère

Ex : PILOTE(NUMPIL,NOMPIL,ADR,SALAIRE) ou AVION(NUMAV,NOMAV,CAP,LOC)

c) *Attributs comparables ou compatibles*

Deux attributs sont comparables ou compatibles si ils sont définis sur le même domaine .

3) *Contraintes d'intégrités relationnelles*

A) *Contraintes d'intégrité statique*

Règles de cohérence de la base induite par les concepts relationnels .

Elles sont invariantes et ne dépendent pas de l'application .

a) *Contrainte de domaine*

Toute valeur affectée à un attribut doit être dans le domaine de l'attribut .

b) *Contrainte de relation*

Toute valeur affectée à la clé primaire doit être unique et obligatoire (doit exister) .

c) *Contrainte de référence*

Toute valeur affectée à la clé étrangère doit exister dans la clé primaire associée .

$Valeur (clé\ étrangère) \subseteq Valeur (clé\ primaire)$

B) *Contraintes dynamiques*

Les règles de cohérence de la BD dépendent de l'application

Ex :

Pour un vol , on a $VILLE_DEP \neq VILLE_ARR$ (\neq : différent) et $HEURE_DEP < HEURE_ARR$

La contrainte statique est vérifiée automatiquement par le SGBD

Les contraintes dynamiques sont spécifiées et programmées par l'administrateur de la BD.

4) Algèbre relationnel

A) Généralités

a) Opérateurs relationnels

Relation unaire

Projection

Selection

Relation binaire

Jointure

Division

b) Opérateurs ensemblistes

Union , qui s'apparente à un OR.

Intersection , qui s'apparente à un AND.

Différence , qui s'apparente à un NOT.

B) Opérateurs relationnels

a) Projection

Déf.: Soit $R(U)$ une relation d'attributs U

$$U = \langle A1, A2, \dots, An \rangle \quad X \subseteq U$$

$$RS = \text{PROJECTION de } R \text{ sur } X$$

$$RS = \text{PROJECTION}(R/X)$$

$$RS = \langle t(X) / t \in R \rangle$$

Exemple :

Quels sont le noms et adresses des pilotes ?

PILOTE	NUMPIL	NOMPIL	ADR	SALAIRE
	100	Dupont	Marseille	5000
	200	Durand	Marseille	4000
	300	Martin	Nice	3000

RS=PROJECTION(PILOTE/NOMPIL,ADR)

RS	NOMPIL	ADR
	Dupont	Marseille
	Durand	Marseille
	Martin	Nice

La projection enlève les doublons

RS=PROJECTION(PILOTE/ADR)

RS	ADR
	Marseille
	Nice

Propriétés :

$$\text{Degré}(RS) \leq \text{Degré}(R)$$

$$\text{Cardinalité}(RS) \leq \text{Cardinalité}(R)$$

b) Sélection

Soient R(U) une relation d'attributs $U = \langle A1, A2, \dots, An \rangle$ constante, une valeur dans le domaine de A

$$RS = SELECTION(R / Ai \theta \text{ constante})$$

$$RS = SELECTION(R / A \text{ Opérateur de comparaison Valeur a comparer})$$

Avec les opérateurs de comparaison suivants :

'='	'<>'
'<'	'<='
'>'	'>='

Exemple :

Chercher les pilotes habitant à Marseille

Avec comme condition ADR='MARSEILLE'

RS=SELECTION(PILOTE/ADR='MARSEILLE')

RS	NUMPIL	NOMPIL	ADR	SALAIRE
	100	Dupont	Marseille	5000
	200	Durand	Marseille	4000

Propriétés :

$$\text{Degré}(RS) = \text{Degré}(R)$$

$$\text{Cardinalité}(RS) \leq \text{Cardinalité}(R)$$

On a ainsi les opérateurs suivants :

La projection : sert à découper verticalement une relation

La selection : sert à découper horizontalement une relation

Toujours commencer par la selection

Exercices :

Quels sont les noms des pilotes habitants à Marseille ?

R1=SELECTION(PILOTE/ADR='Marseille')

RS=PROJECTION(R1/NOMPIL)

RS	NOMPIL
	Dupont
	Durand

c) Jointure

La jointure est un opérateur binaire , elle nécessite deux relations

Soient R(U1), S(U2) deux relations d'attributs

$$U1 = \langle A1, A2, \dots, An \rangle$$

$$\text{et } U2 = \langle B1, B2, \dots, Bn \rangle$$

et $Ai \in U1$ et $Bj \in U2$ deux attributs compatibles

$$RS = \text{JOINTURE}(R, S / Ai \theta Bj)$$

$$RS = \langle (t, s) / t \in R, s \in S, t(Ai) \theta S(Bj) \rangle$$

avec θ : un opérateur de comparaison

Le schéma de RS contient tous les attributs de A et tous les attributs de S

R	A	B	C	S	A	D
	a1	b1	c1		a1	d1
	a2	b1	c1		a1	d2
	a3	b2	c1		a2	d3
					a2	d4

RS=JOINTURE(R,S/As=Ar)

RS	A	B	C	A	D
	a1	b1	c1	a1	d1
	a1	b1	c1	a1	d2
	a2	b1	c1	a2	d3
	a2	b1	c1	a2	d4

Quels sont les noms des pilotes en service ?

PILOTE	NUMPIL	NOMPIL	ADR	SAL	VOL	NUMVOL	NUMPIL#	NUMAV#	VILLE_DEP	VILLE_ARR
	10	Dupont	Nice	5000	IT100	10	100		Marseille	Paris
	20	Durand	Marseille	6000	IT200	10	100		Paris	Marseille
	30	Dupont	Marseille	4000	IT300	30	200		Toulouse	Paris

R1=JOINTURE(PILOTE,VOL/NUMPIL=NUMPIL)

RS=PROJECTION(R1/NOMPIL)

R1	NUMPIL	NOMPIL	ADR	SAL	NUMVOL	NUMPIL	NUMAV	VILLE_DEP	VILLE_ARR	RS	NOMPIL
	10	Dupont	Nice	5000	IT100	10	100	Marseille	Paris		Dupont
	10	Dupont	Nice	5000	IT200	10	100	Paris	Marseille		Dupont
	30	Dupont	Marseille	4000	IT300	30	200	Toulouse	Paris		

Différents types de jointures :

L'équijointure

RS=JOINTURE(PILOTE,VOL/NUMPIL=NUMPIL)

Jointure naturelle

Enlève la deuxième occurrence d'un attribut identique , ici , enlève NUMPIL

RS	NUMPIL	NOMPIL	ADR	SAL	NUMVOL	NUMAV	VILLE_DEP	VILLE_ARR
	10	Dupont	Nice	5000	IT100	100	Marseille	Paris
	10	Dupont	Nice	5000	IT200	100	Paris	Marseille
	30	Dupont	Marseille	4000	IT300	200	Toulouse	Paris

Jointure par <>

RS=JOINTURE(PILOTE,VOL/NUMPIL<>NUMPIL)

RS	NUMPIL	NOMPIL	ADR	SAL	NUMVOL	NUMPIL	NUMAV	VILLE_DEP	VILLE_ARR
	10	Dupont	Nice	5000	IT300	30	200	Toulouse	Paris
	20	Durand	Marseille	6000	IT100	10	100	Marseille	Paris
	20	Durand	Marseille	6000	IT200	10	100	Paris	Marseille
	20	Durand	Marseille	6000	IT300	30	200	Toulouse	Paris
	30	Dupont	Marseille	4000	IT100	10	100	Marseille	Paris
	30	Dupont	Marseille	4000	IT200	10	100	Paris	Marseille

Autojointure

RS=JOINTURE(PILOTE,PILOTE/NUMPIL,NUMPIL)

RS	NUMPIL	NOMPIL	ADR	SAL	NUMPIL	NOMPIL	ADR	SAL
	10	Dupont	Nice	5000	10	Dupont	Nice	5000
	20	Durand	Marseille	6000	20	Durand	Marseille	6000
	30	Dupont	Marseille	4000	30	Dupont	Marseille	4000

d) La division

Soient R(B,A) et S(A) deux relations respectivement une relation binaire et une unaire avec Ar et As définis sur le même domaine .

RS=DIVISION(R,S/Ar,As)

Ar et As n'ont pas forcément le même nom .

$$RS = \langle t(B) \mid t \in R \text{ et } \forall s \in S, (t(B), S(A)) \in R \rangle$$

Contraintes de la division :

R a deux attributs , S un seul.

Les attributs de la division doivent être compatibles .

RS=DIVISION(R,S/A,A)

R	B	A	S	A	RS	B
	b1	a1		a1		b1
	b1	a2		a2		
	b2	a1				

b1 est associé dans R à toutes les valeurs de A.

$$B = \{b1, b2\} \Rightarrow (b1, a1), (b1, a2)$$

Or b2 n'est pas associé dans R à a2 , donc il n'est pas pris

La division exprime “tous les” , ou bien “au moins tous les” $\Rightarrow \forall$

La jointure exprime “un” , “il existe” $\Rightarrow \exists$

Exemple : numéros de pilotes qui conduisent tous les avions de sa compagnie

NUMAV est l'attribut de la division

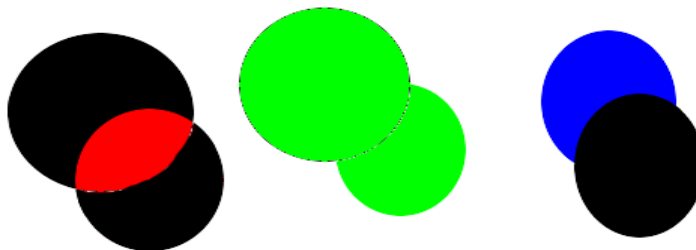
R1=PROJECTION(AVION/NUMAV)

R2=PROJECTION(VOL/NUMAV,NOMPIL)

RS=DIVISION(R1,R2/NUMAV,NUMAV)

Les pilotes apparaissant dans le résultat sont associés à tous les avions de la compagnie.

C) Opérateurs ensemblistes



ROUGE:

INTERSECTION $R \cap S$

VERT : UNION $R \cup S$

BLEU : DIFFERENCE $R - S$

$RS = \text{INTERSECTION}(R, S) = \{t/t \in R \text{ et } t \in S\}$

$RS = \text{UNION}(R, S) = \{t/t \in R \text{ ou } t \in S\}$

$RS = \text{DIFFERENCE}(R, S) = \{t/t \in R \text{ et } t \notin S\}$

R et S sont uni-compatibles si

-Degré(R)=Degré(S)

-les attributs de R et de S sont définis sur le même domaine

R(A, B, C) A même domaine que D

S(D, E, F) B même domaine que E

C même domaine que F

Exemple : Noms des pilotes habitant Nice et gagnant 5000€

$R1 = \text{SELECTION}(\text{PILOTE}/\text{ADR}='Nice')$

$R2 = \text{SELECTION}(\text{PILOTE}/\text{SAL}=5000)$

$R3 = \text{INTERSECTION}(R1, R2)$

$RS = \text{PROJECTION}(R3/\text{NOMPIL})$

$R1 = \text{SELECTION}(\text{PILOTE}/\text{ADR}='Nice')$

$R2 = \text{SELECTION}(R1/\text{SALAIRE}=5000)$

$RS = \text{PROJECTION}(R2/\text{NOMPIL})$

Numéros des pilotes qui n'ont jamais assurés de vol

$R1 = \text{PROJECTION}(\text{VOL}/\text{NUMPIL})$

$R2 = \text{PROJECTION}(\text{PILOTE}/\text{NUMPIL})$

$R3 = \text{DIFFERENCE}(R2, R1)$

$R4 = \text{JOINTURE}(\text{PILOTE}, R3/\text{NUMPIL}=\text{NUMPIL})$

$RS = \text{PROJECTION}(R4/\text{NOMPIL})$

III. DEPENDANCES FONCTIONNELLES

1) Notations

$r(\mathbb{R})$ est une relation r d'attribut \mathbb{R}

\mathbb{R} est un ensemble d'attributs

$\mathbb{R} = \{A, B, C, D\} = ABCD$

$X \subseteq \mathbb{R}$ X est un ensemble de \mathbb{R}

$X = \{A, B\} = AB$

$X \in P(\mathbb{R})$ ensemble des parties

Exemple :

$P(\mathbb{R})$: L'ensemble de parties de \mathbb{R}

Tous ces sous-ensembles sont :

$\mathbb{R} = \{A, B, C\}$

$P(\mathbb{R}) = \{\emptyset, A, B, C, AB, AC, BC, ABC\}$

$P(\mathbb{R}) = \{\{\emptyset\}, \{A\}, \dots, \{A, B, C\}\} = 2^{\mathbb{R}}$

$|P(\mathbb{R})| = \text{card}(P(\mathbb{R})) = 2^{\mathbb{R}} = 2^{\text{card}(\mathbb{R})}$

2) Définitions

Soit $r(\mathbb{R})$ une relation d'attributs \mathbb{R} , X, Y deux sous ensembles de \mathbb{R} ($X, Y \subseteq \mathbb{R}$)

On dit que Y dépend fonctionnellement de X ou X détermine Y ; $X \Rightarrow Y$ si et seulement si

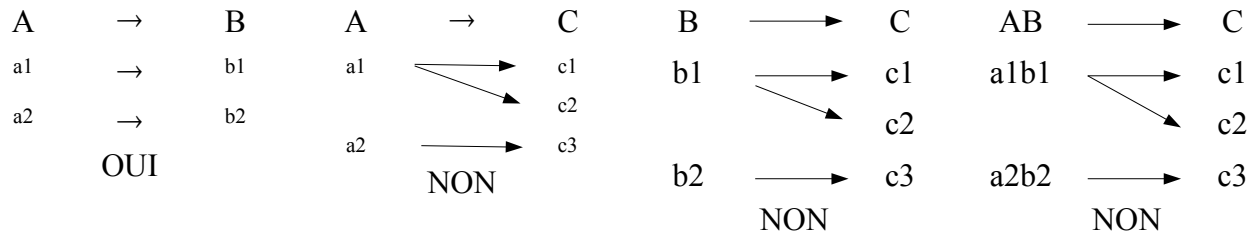
$\forall t, t' \in r$ si $t(x) = t'(x)$

Alors $t(y) = t'(y)$

r	A	B	C
t1	a1	b1	c1
t2	a1	b1	c2
t3	a2	b2	c3

$A \rightarrow B$ car t1 t2 t3 idems
 $A \rightarrow C$ NON
 $C \rightarrow B$ car il n'existe pas de couple de tuples
 $B \rightarrow C$ NON
 $AB \rightarrow C$ NON
 $C \rightarrow AB$

Retrouver une dépendance fonctionnelle en mode graphique :



Méthode par comptage (projeter , compter)

$$x \rightarrow y$$

$$|r(X)| = |r(xy)|$$

$$|r(A)| = |r(AB)|$$

$$2 = 2$$

Donc $A \rightarrow AB$
 $A \rightarrow B$

$$|r(A)| = |r(AC)|$$

$$2 \neq 3$$

Donc A n'implique pas AC ,

et A n'implique pas C

Les dépendances fonctionnelles sont des contraintes d'intégrités qui devraient être vérifiées quelque soit le tuple de la relation .

3) Axiomes d'Amstrong

Soient w, x, y, z des sous-ensembles d'attributs de \mathbb{R} on a les propriétés suivantes :

a) La réflexivité

Si $y \subseteq x$ alors $x \Rightarrow y$ ($x \Rightarrow x$)

ex:

$$x = \{A, B, C\} \quad y = \{A, B\}$$

$$AB \subseteq ABC$$

$$ABC \Rightarrow AB$$

b) L'augmentation

Si $x \Rightarrow y$ et $z \subseteq w$ alors $xw \Rightarrow yz$

ex:

$$A \Rightarrow B$$

$$ACD \Rightarrow BC$$

$$AC \Rightarrow B \emptyset$$

car $\emptyset \subseteq C$

c) La transitivité

Si $x \Rightarrow y$ et $y \Rightarrow z$ alors $x \Rightarrow z$

d) La pseudo-transitivité

Si $x \Rightarrow y$ et $yz \Rightarrow t$ alors $xz \Rightarrow t$

e) L'union

Si $x \Rightarrow y$ et $x \Rightarrow z$ alors $x \Rightarrow yz$

f) La décomposition

Si $x \Rightarrow yz$ alors $x \Rightarrow y$ et $x \Rightarrow z$

4) Autres concepts

a) Dépendance fonctionnelle minimale

$x \Rightarrow y$ Est une DF minimale valide dans $r(\mathbb{R})$ si et seulement si $x \Rightarrow y$ avec $|Y|=1$

$\forall x \in X, X-x$ n'implique pas y

Tous les attributs de X sont nécessaire pour déterminer Y

Si $A \rightarrow B$
 $AC \rightarrow B$ n'est pas minimale car C non nécessaire

b) Clé candidate

Soit $r(\mathbb{R})$ une relation d'attribut $\mathbb{R} \quad x \subseteq \mathbb{R}$

x est une clé candidate si et seulement si

- $x \rightarrow \mathbb{R}$
- $|r(x)| = |r(\mathbb{R})|$

exemple :

$C \rightarrow ABC$
 $AC \rightarrow ABC$
 $ABC \rightarrow ABC$
 $BC \rightarrow ABC$

r	A	B	C
	a1	b1	c1
	a1	b1	c2
	a2	b2	c3

c) Clé candidate minimale

Soit $r(\mathbb{R})$ une relation d'attributs \mathbb{R} tel que $X \subseteq \mathbb{R}$

X est clé candidate minimale si et seulement si :

- $X \Rightarrow \mathbb{R}$ X clé candidate
- $\forall x \in x, X - x$, n'implique pas \mathbb{R} (contrainte de minimalité)

ex:

$C \rightarrow ABC$

- C est clé candidate
- C est une clé candidate minimale

$AC \rightarrow ABC$

- AC clé candidate
- AC n'est pas une clé candidate minimale

d) Clé primaire

Soit $r(\mathbb{R})$ une relation d'attribut \mathbb{R}

$$X \subseteq \mathbb{R}$$

X est une clé primaire si et seulement si X est une clé candidate minimale.

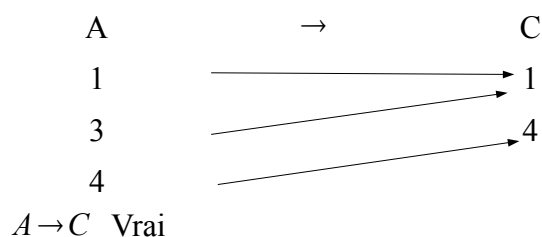
Choix d'une clé primaire à partir de plusieurs clé candidates minimales

X,Y deux clés candidates minimales X est choisi si :

- $|X| \leq |Y|$
- Regarder les types $|X|=|Y|$ numériques prioritaires

r	A	B	C	D	E
	1	1	1	1	1
	1	1	1	2	1
	3	1	1	3	1
	4	1	4	4	1
	1	5	1	1	5

$$\mathbb{R} = \{A, B, C, D, E\}$$



e) Attributs non-clé primaire / non clé candidate minimale

Un attribut A est dit non clé primaire s'il n'appartient pas à la clé primaire .

Ex: R1(A,B,C,D) et R2(A,C,D)

C et D sont les attributs non clé primaires dans R1

Il y a deux clé candidates minimales dans R1 : A et B

Dans R2 , C et D sont des attributs non clé primaire

IV. NORMALISATION DES RELATIONS

1) Objectif

Conception d'un bon schéma relationnel

- Redondance des données
- Anomalies de stockage
 - Anomalie d'ajout
 - Anomalie de modification
 - Anomalie de suppression

ENSEIGNANTS	NOM	ADR	FONCT	SAL
	A	Marseille	Prof	10000
	B	Nice	Prof	10000
	C	Toulouse	Prof	10000
	D	Paris	Assistant	2000
	E	Paris	Mc	5000
	F	Paris	Mc	5000
	G	Aix	Mc	5000

Redondance de données : (Prof,10000)

Anomalie d'ajout :

Impossible d'ajouter une nouvelle fonction et un nouveau salaire

Anomalie de modification :

Liée à la redondance des données

Si on augmente le salaire des Prof de 10% , plusieurs opérations de modifications ont lieu pour les même valeurs

Anomalie de suppression

Si on supprime la fonction Assistant , on perd définitivement le salaire des assistants

Cause : DF entre FONCT et SAL

Forme normale

L'objectif est la formation d'un bon schéma relationnel : moins d'anomalies de stockages et de redondance des données

2) Formes normales

1NF, 2NF, 3NF, BCNF constituent les différents degrés de normalité d'une relation.

Plus le degré est important, moins on a de redondance des données et d'anomalies de stockages

A) 1NF

Une relation $r(\mathbb{R})$ est en forme normale si tous ses attributs sont monovalués (à valeur simple, unique)

Ex :

ENSEIGNANT est en 1NF

Monovalué : pour chaque tuple, la valeur d'un attribut est simple (n'est pas un ensemble)

Exemple de relation avec un attribut multivalué :

EMPLOYE	<u>NOM</u>	ADR	TEL
	Dupont	Marseille	{t1,t2,t3}
	Durand	Nice	{t4,t5}
	Martin	Aix	{t6}

Passage de N1NF à une forme 1NF (normalisation en 1NF)

Deux cas :

Attribut multivalué à cardinalité fixe :

EMPLOYE	<u>NOM</u>	ADR	TEL1	TEL2	TEL3
	Dupont	Marseille	t1	t2	t3
	Durand	Nice	t4	t5	
	Martin	Aix	t6		

Cardinalité variable (ajout de relations toutes clés)

Les attributs de la relation forment la clé primaire

EMPLOYE	<u>NOM</u>	ADR	TELEPHONE	<u>NOM</u>	<u>TEL</u>
	Dupont	Marseille		Dupont	t1
	Durand	Nice		Dupont	t2
	Martin	Aix		Dupont	t3
				Durand	t4
				Durand	t5
				Martin	t6

B) 2NF

Une relation $r(\mathbb{R})$ est en 2NF si et seulement si

Elle est en 1NF

N'a pas de DF entre une partie de la clé primaire et un attribut non clé primaire

Ex:

$r(\underline{A}, B, C, D) \quad AB \rightarrow ABCD$

$r1(\underline{A}, B, C, D) \quad A \rightarrow ABCD \quad 2NF$

$r2(\underline{A}, B, C, D) \quad \begin{matrix} AB \rightarrow ABCD \\ B \rightarrow C \end{matrix} \quad \text{pas 2NF (pour } B \rightarrow C \text{)}$

Normalisation en 2NF

Théorème de décomposition :

Soit $r(X, Y, Z)$ une relation

$X \rightarrow Y$

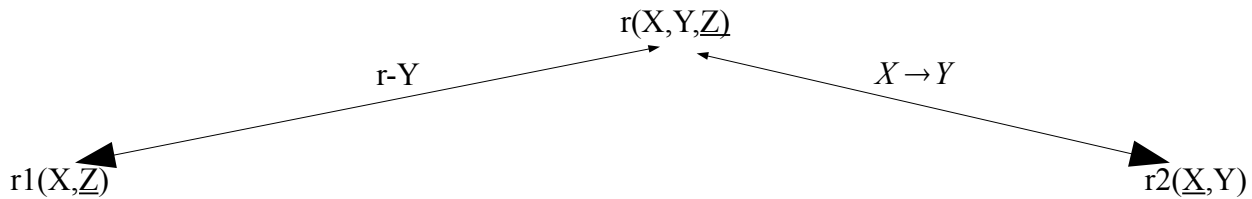
$r = \text{JOINTURE_NATURELLE}(r1, r2 / X=X)$

avec

$r1 = \text{PROJECTION}(r / X, Z)$

$r2 = \text{PROJECTION}(r / X, Y)$

La décomposition de r en $r1$ et $r2$ est sans perte d'informations (ni tuples, ni attributs)



Aucune perte d'informations

C) 3NF

Une relation $r(\mathbb{R})$ est en 3NF si et seulement si

Elle est en 2NF

Pas de DF entre attributs non clés candidates minimales

$r(\underline{A}, \underline{B}, C, D)$

Clé candidate minimale AB (qui est aussi clé primaire)

$AB \rightarrow ABCD$

$C \rightarrow D$

r est en 2NF , pas en 3NF car $C \rightarrow D$

Ex :

PRODUIT	NO_PROD	CODE_TVA	TAUX_TVA
	P1	1	18.6
	P2	1	18.6
	P3	2	33
	P4	1	18.6
	P5	1	18.6

$NO_PROD \rightarrow CODE_TVA, TAUX_TVA$

$CODE_TVA \rightarrow TAUX_TVA$

PRODUIT est en 1NF , car les valeurs des attributs sont simples

PRODUIT est en 2NF , car la clé primaire comporte un attribut

PRODUIT n'est pas en 3NF , car $CODE_TVA \rightarrow TAUX_TVA$

Anomalie de suppression : si on supprime P3 , on perd la fonction (-,2,33)

Anomalie de redondance : (-,1,18.6) ; Anomalie d'ajout : (-,3,15) serait impossible

TVA	CODE_TVA	TAUX_TVA
	1	18.6
	2	33

PRODUIT	NO_PROD	CODE_TVA
	P1	1
	P2	1
	P3	2
	P4	1
	P5	1