

2. CONCEPTS DES BASES DE DONNEES

Hend AFFES
Jean-Yves Tigli

BAT3 – Polytech Nice – Sophia

2013-2014

2. CONCEPTS DES BASES DE DONNEES

Contenu

1. **Tables, lignes et colonnes**
2. **Valeur *null***
3. **Identifiants et clés étrangères**
4. **Schéma et contenu**
5. **Exemple de base de données**
6. **Modifications et contraintes**
7. **Redondances internes**

2.1 Tables, lignes et colonnes

CLIENT					
NCLI	NOM	ADRESSE	LOCALITE	(CAT)	COMPTE
B062	GOFFIN	72, r. de la Gare	Namur	B2	-3200
B112	HANSENNE	23, r. Dumont	Poitiers	C1	1250
B332	MONTI	112, r. Neuve	Genève	B2	0
B512	GILLET	14, r. de l'Eté	Toulouse	B1	-8700
C003	AVRON	8, r. de la Cure	Toulouse	B1	-1700
C123	MERCIER	25, r. Lemaître	Namur	C1	-2300
C400	FERARD	65, r. du Tertre	Poitiers	B2	350
D063	MERCIER	201, bvd du Nord	Toulouse		-2250
F010	TOUSSAINT	5, r. Godefroid	Poitiers	C1	0
F011	PONCELET	17, Clos des Erables	Toulouse	B2	0
F400	JACOB	78, ch. du Moulin	Bruxelles	C2	0
K111	VANBIST	180, r. Florimont	Lille	B1	720
K729	NEUMAN	40, r. Bransart	Toulouse		0
L422	FRANCK	60, r. de Wépion	Namur	C1	0
S127	VANDERKA	3, av. des Roses	Namur	C1	-4580
S712	GUILLAUME	14a, ch. des Roses	Paris	B1	0

2.1 Tables, lignes et colonnes

CLIENT					
NCLI	NOM	ADRESSE	LOCALITE	(CAT)	COMPTE
B062	GOFFIN	72, r. de la Gare	Namur	B2	-3200
B112	HANSENNE	23, r. Dumont	Poitiers	C1	1250
B332	MONTI	112, r. Neuve	Genève	B2	0
B512	GILLET	14, r. de l'Eté	Toulouse	B1	-8700
C003	AVRON	8, r. de la Cure	Toulouse	B1	-1700
C123	MERCIER	25, r. Lemaître	Namur	C1	-2300
C400	FERARD	65, r. du Tertre	Poitiers	B2	350
D063	MERCIER	201, bvd du Nord	Toulouse		-2250
F010	TOUSSAINT	5, r. Godefroid	Poitiers	C1	0
F011	PONCELET	17, Clos des Erables	Toulouse	B2	0
F400	JACOB	78, ch. du Moulin	Bruxelles	C2	0
K111	VANBIST	180, r. Florimont	Lille	B1	720
K729	NEUMAN	40, r. Bransart	Toulouse		0
L422	FRANCK	60, r. de Wépion	Namur	C1	0
S127	VANDERKA	3, av. des Roses	Namur	C1	-4580
S712	GUILLAUME	14a, ch. des Roses	Paris	B1	0

schéma (blue bracket on the right side of the table)
données (red bracket on the right side of the table)
ligne (red bracket on the left side of the row B512)
colonne obligatoire (blue bracket below the columns NOM, ADRESSE, LOCALITE)
colonne facultative (blue bracket below the columns (CAT), COMPTE)

2.2 Valeur *null*

L'absence de valeur est indiquée par un marqueur spécial, dit *valeur null*. Généralement représenté par `<null>` ou par *rien*.

Problème : plusieurs interprétations possibles

1. information pertinente mais inexistante pour l'entité
2. information non pertinente pour cette entité
3. information existante mais actuellement inconnue

Recommandation : éviter si possible les colonnes facultatives.

2.3 Identifiants et clés étrangères

Un **identifiant** est un groupe de colonnes d'une table T tel qu'il ne puisse, à tout moment, exister plus d'une ligne dans T qui possède des valeurs déterminées pour ces colonnes. La valeur de l'identifiant permet de désigner une ligne de T .

Une **clé étrangère** est un groupe de colonnes d'une table S tel qu'il existe, à tout moment, dans une table T , une ligne dont l'identifiant a pour valeur(s) celle(s) de ce groupe. La valeur de la clé étrangère sert à référencer une ligne de la table T .

2.3 Identifiants et clés étrangères

CLIENT					
NCLI	NOM	ADRESSE	LOCALITE	(CAT)	COMPTE
B062	GOFFIN	72, r. de la Gare	Namur	B2	-3200
B112	HANSENNE	23, r. Dumont	Poitiers	C1	1250
B332	MONTI	112, r. Neuve	Genève	B2	0
B512	GILLET	14, r. de l'Été	Toulouse	B1	-8700
C003	AVRON	8, r. de la Cure	Toulouse	B1	-1700
C123	MERCIER	25, r. Lemaître	Namur	C1	-2300
C400	FERARD	65, r. du Tertre	Poitiers	B2	350
D063	MERCIER	201, bvd du Nord	Toulouse		-2250
F010	TOUSSAINT	5, r. Godefroid	Poitiers	C1	0
F011	PONCELET	17, Clos des Erables	Toulouse	B2	0
F400	JACOB	78, ch. du Moulin	Bruxelles	C2	0
K111	VANBIST	180, r. Florimont	Lille	B1	720
K729	NEUMAN	40, r. Bransart	Toulouse		0
L422	FRANCK	60, r. de Wépion	Namur	C1	0
S127	VANDERKA	3, av. des Roses	Namur	C1	-4580
S712	GUILLAUME	14a, ch. des Roses	Paris	B1	0

identifiant

COMMANDE		
NCOM	NCLI	DATECOM
30178	K111	21/12/2008
30179	C400	22/12/2008
30182	S127	23/12/2008
30184	C400	23/12/2008
30185	F011	2/01/2009
30186	C400	2/01/2009
30188	B512	3/01/2009

clé étrangère

dans le schéma
dans les données

2.3 Identifiants et clés étrangères

Un **identifiant** définit une **contrainte d'unicité**.

Une table peut posséder plusieurs identifiants. On choisit l'un d'eux, qu'on déclare **primaire**. Les autres sont dès lors **secondaires**.

L'identifiant primaire est constitué de **colonnes obligatoires**.

Un identifiant est **minimal** si chacune de ses colonnes est nécessaire pour garantir la contrainte d'unicité.

Il est possible de déclarer une **table sans identifiant** mais ceci n'est pas recommandé.

2.3 Identifiants et clés étrangères

Une clé étrangère définit une **contrainte référentielle**.

Une clé étrangère référence l'**identifiant primaire** de la table cible.

Une clé étrangère et l'identifiant qu'elle référence ont la **même composition** : même nombre de colonnes et colonnes de mêmes types prises deux à deux.

Clé primaire = *primary key*.

Clé étrangère = *foreign key*.

2.4 Schéma et contenu

CLIENT					
NCLI	NOM	ADRESSE	LOCALITE	(CAT)	COMPTE
B062	GOFFIN	72, r. de la Gare	Namur	B2	-3200
B112	HANSENNE	23, r. Dumont	Poitiers	C1	1250
B332	MONTI	112, r. Neuve	Genève	B2	0
B512	GILLET	14, r. de l'Eté	Toulouse	B1	-8700
C003	AVRON	8, r. de la Cure	Toulouse	B1	-1700
C123	MERCIER	25, r. Lemaître	Namur	C1	-2300
C400	FERARD	65, r. du Tertre	Poitiers	B2	350
D063	MERCIER	201, bvd du Nord	Toulouse		-2250
F010	TOUSSAINT	5, r. Godefroid	Poitiers	C1	0
F011	PONCELET	17, Clos des Erables	Toulouse	B2	0
F400	JACOB	78, ch. du Moulin	Bruxelles	C2	0
K111	VANBIST	180, r. Florimont	Lille	B1	720
K729	NEUMAN	40, r. Bransart	Toulouse		0
L422	FRANCK	60, r. de Wépion	Namur	C1	0
S127	VANDERKA	3, av. des Roses	Namur	C1	-4580
S712	GUILLAUME	14a, ch. des Roses	Paris	B1	0

le schéma

les données

2.4 Schéma et contenu

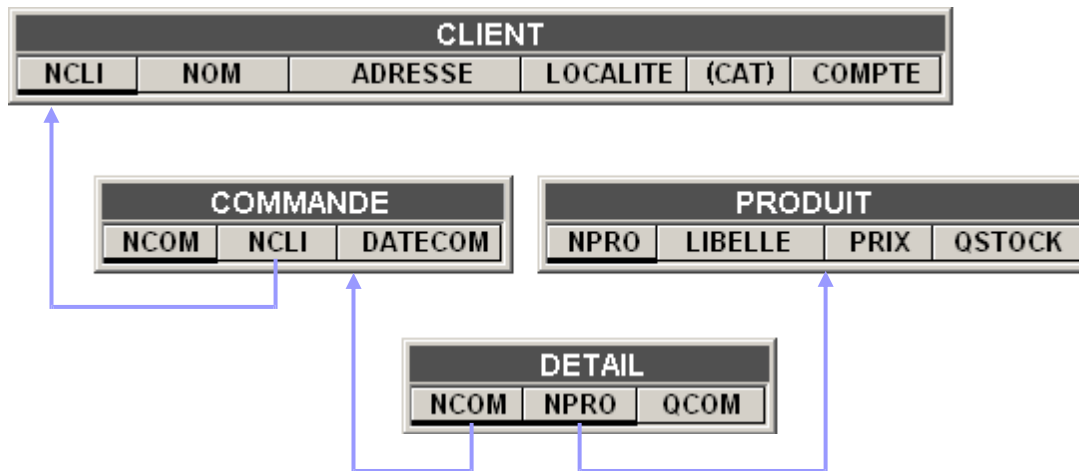
Le **schéma d'une table** définit sa structure. Il spécifie notamment :

1. le nom de la table,
2. pour chaque colonne, son nom, son type, son caractère obligatoire,
3. l'identifiant primaire (liste des colonnes)
4. les identifiants secondaires éventuels (liste des colonnes)
5. les clés étrangères éventuelles (liste des colonnes et table cible).

Le **contenu d'une table** est formé d'un ensemble de lignes conformes au schéma.

Le **contenu d'une table** est sujet à de fréquentes modifications. Le **schéma d'une table** peut évoluer mais moins fréquemment.

2.5 Exemple de base de données



2.5 Exemple de base de données

CLIENT					
NCLI	NOM	ADRESSE	LOCALITE	(CAT)	COMPTE
B062	GOFFIN	72, r. de la Gare	Namur	B2	-3200
B112	HANSENNE	23, r. Dumont	Poitiers	C1	1250
B332	MONTI	112, r. Neuve	Genève	B2	0
B512	GILLET	14, r. de l'Eté	Toulouse	B1	-8700
C003	AVRON	8, r. de la Cure	Toulouse	B1	-1700
C123	MERCIER	25, r. Lemaître	Namur	C1	-2300
C400	FERARD	65, r. du Tertre	Poitiers	B2	350
D063	MERCIER	201, bvd du Nord	Toulouse		-2250
F010	TOUSSAINT	5, r. Godefroid	Poitiers	C1	0
F011	PONCELET	17, Clos des Erables	Toulouse	B2	0
F400	JACOB	78, ch. du Moulin	Bruxelles	C2	0
K111	VANBIST	180, r. Florimont	Lille	B1	720
K729	NEUMAN	40, r. Bransart	Toulouse		0
L422	FRANCK	60, r. de Wépion	Namur	C1	0
S127	VANDERKA	3, av. des Roses	Namur	C1	-4580
S712	GUILLAUME	14a, ch. des Roses	Paris	B1	0

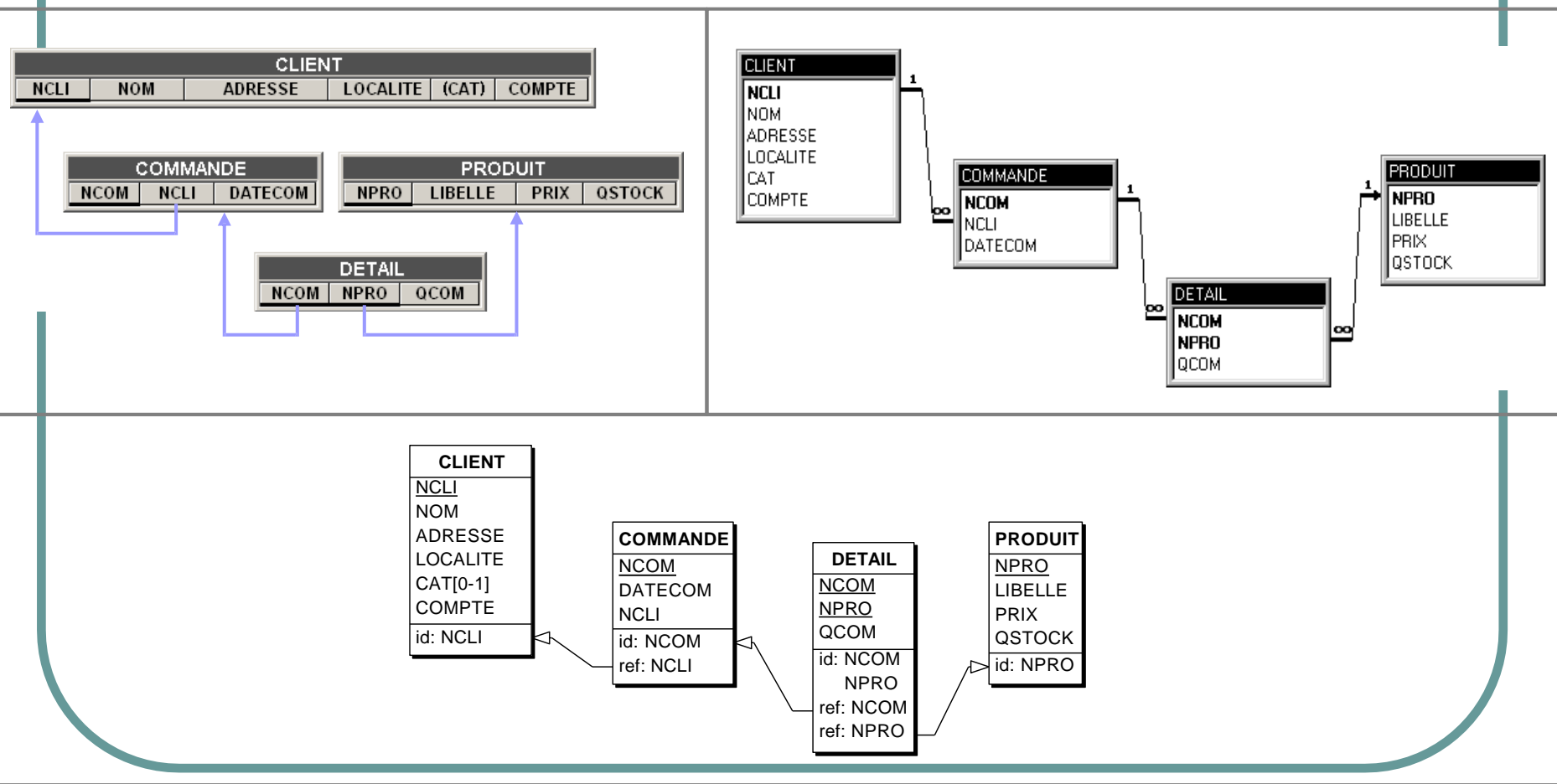
PRODUIT			
NPRO	LIBELLE	PRIX	QSTOCK
CS262	CHEV. SAPIN 200x6x2	75	45
CS264	CHEV. SAPIN 200x6x4	120	2690
CS464	CHEV. SAPIN 400x6x4	220	450
PA45	POINTE ACIER 45 (2K)	105	580
PA60	POINTE ACIER 60 (1K)	95	134
PH222	PL. HETRE 200x20x2	230	782
PS222	PL. SAPIN 200x20x2	185	1220

COMMANDE		
NCOM	NCLI	DATECOM
30178	K111	21/12/2008
30179	C400	22/12/2008
30182	S127	23/12/2008
30184	C400	23/12/2008
30185	F011	2/01/2009
30186	C400	2/01/2009
30188	B512	3/01/2009

DETAIL		
NCOM	NPRO	QCOM
30178	CS464	25
30179	CS262	60
30179	PA60	20
30182	PA60	30
30184	CS464	120
30184	PA45	20
30185	CS464	260
30185	PA60	15
30185	PS222	600
30186	PA45	3
30188	CS464	180
30188	PA45	22
30188	PA60	70
30188	PH222	92

2.5 Exemple de base de données

Variantes de schéma



2.6 Modifications et contraintes d'intégrité

On admet trois contraintes d'intégrité :

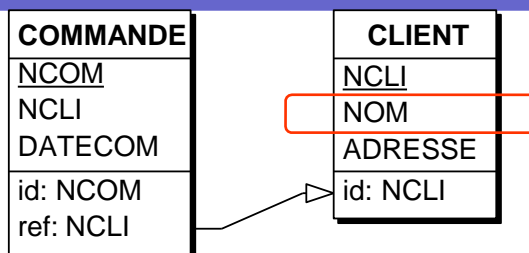
1. colonne obligatoire
2. contrainte d'unicité
3. contrainte référentielle.

Il existe trois opérations élémentaires de modification :

1. insérer une ligne
2. supprimer une ligne
3. modifier une valeur de colonne d'une ligne.

Principe : une modification ne sera effectuée que si son résultat ne viole aucune contrainte d'intégrité.

2.6 Modifications et contraintes - Colonne obligatoire



colonne NOM obligatoire

insérer une ligne
de CLIENT

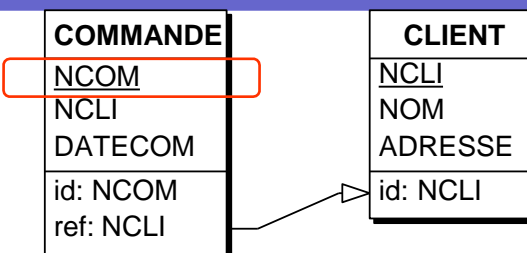
la valeur de NOM doit être non *null*

supprimer une ligne
de CLIENT

modifier valeur de NOM
de CLIENT

la nouvelle valeur de NOM doit être non *null*

2.6 Modifications et contraintes - Identifiant



{NCOM} identifiant de COMMANDE

**insérer une ligne
de COMMANDE**

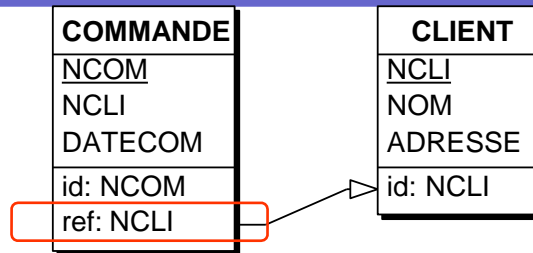
**la valeur de NCOM ne doit pas être déjà
présente dans une ligne de COMMANDE**

**supprimer une ligne
de COMMANDE**

**modifier valeur de NCOM
de COMMANDE**

**la nouvelle valeur de NCOM ne doit pas être
déjà présente dans une ligne de COMMANDE**

2.6 Modifications et contraintes - Clé étrangère



{NCLI} clé étrangère de COMMANDE
vers CLIENT

insérer une ligne
de COMMANDE

la valeur de NCLI doit être présente dans la
colonne NCLI d'une ligne de CLIENT

supprimer une ligne
de COMMANDE

supprimer une ligne
de CLIENT

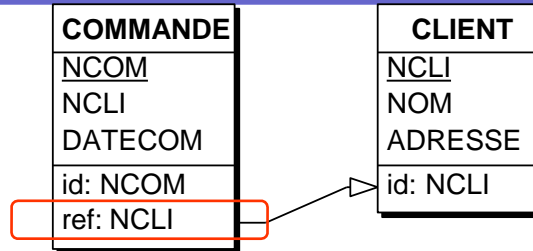
**l'intégrité référentielle doit être satisfaite
après l'opération**

modifier valeur de NCLI
de COMMANDE

la nouvelle valeur de NCLI doit être présente
dans la colonne NCLI d'une ligne de CLIENT

**plusieurs
comportements
possibles**

2.6 Modifications et contraintes - Clé étrangère



	Comportement	{NCLI} clé étrangère de COMMANDE vers CLIENT
supprimer une ligne de CLIENT	1. mode <i>no action</i>	opération refusée
	2. mode <i>cascade</i>	ligne supprimée mais aussi les lignes de COMMANDE dépendantes
	3. mode <i>set null</i>	(si NCLI de COMMANDE facultative) la colonne NCLI des lignes dépendantes de COMMANDE est mise à <i>null</i>
	4. mode <i>set default</i>	(si <i>default</i> pour NCLI de COMMANDE) la colonne NCLI des lignes dépendantes de COMMANDE est mises à <i>la valeur par défaut</i>

2.7 Redondances internes

Table classant les livres d'une bibliothèque :

LIVRE					
NUMERO	TITRE	AUTEUR	ISBN	DATE_ACHAT	EMPL
1029	L'humanité perdue	Finkelkraut A.	2 02 033300 7	14/10/2008	F3
1030	L'humanité perdue	Finkelkraut A.	2 02 033300 7	14/10/2008	F3
1032	Mercure	Nothomb A.	2 253 14911 X	14/10/2008	G5
1045	Eva Luna	Allende I.	2 253 05354 6	22/2/2009	F3
1067	Mercure	Nothomb A.	2 253 14911 X	24/2/2009	G5
1022	Mercure	Nothomb A.	2 253 14911 X	3/10/2008	G6

2.7 Redondances internes

Observation

Les données TITRE et AUTEUR sont répétées autant de fois qu'il existe de livres identiques.

LIVRE					
NUMERO	TITRE	AUTEUR	ISBN	DATE_ACHAT	EMPL
1029	L'humanité perdue	Finkelkraut A.	2 02 033300 7	14/10/2008	F3
1030			2 02 033300 7	14/10/2008	F3
1032	Mercure	Nothomb A.	2 253 14911 X	14/10/2008	G5
1045	Eva Luna	Allende I.	2 253 05354 6	22/2/2009	F3
1067			2 253 14911 X	24/2/2009	G5
1022			2 253 14911 X	3/10/2008	G6

Cette table viole le principe des bases de données : ***tout fait du domaine d'application est enregistré une et une seule fois.***

2.7 Redondances internes

LIVRE					
NUMERO	TITRE	AUTEUR	ISBN	DATE_ACHAT	EMPL
1029	L'humanité perdue	Finkelkraut A.	2 02 033300 7	14/10/2008	F3
1030	L'humanité perdue	Finkelkraut A.	2 02 033300 7	14/10/2008	F3
1032	Mercure	Nothomb A.	2 253 14911 X	14/10/2008	G5
1045	Eva Luna	Allende I.	2 253 05354 6	22/2/2009	F3
1067	Mercure	Nothomb A.	2 253 14911 X	24/2/2009	G5
1022	Mercure	Nothomb A.	2 253 14911 X	3/10/2008	G6


Problèmes

- gaspillage d'espace
- si on modifie la valeur d'un titre, il faut répercuter cette modification dans toutes les lignes similaires
- si on supprime l'unique exemplaire d'un livre, on perd les informations sur son auteur et son titre
- est-on certain que le titre et l'auteur ont été orthographiés exactement de la même manière pour tous les exemplaires d'un livre ?

2.7 Redondances internes

Suggestion

Rassembler les données communes (ISBN, TITRE, AUTEUR) dans une table spécifique



OUVRAGE		
ISBN	TITRE	AUTEUR
2 02 033300 7	L'humanité perdue	Finkelkraut A.
2 253 14911 X	Mercure	Nothomb A.
2 253 05354 6	Eva Luna	Allende I.

EXEMPLAIRE			
NUMERO	ISBN	DATE_ACHAT	EMPL
1029	2 02 033300 7	14/10/2008	F3
1030	2 02 033300 7	14/10/2008	F3
1032	2 253 14911 X	14/10/2008	G5
1045	2 253 05354 6	22/2/2009	F3
1067	2 253 14911 X	24/2/2009	G5
1022	2 253 14911 X	3/10/2008	G6

2.7 Redondances internes

Deux questions

1. Comment détecter les situations de redondance ?
2. Comment les corriger ?

La réponse à ces questions repose sur une nouvelle forme de contrainte d'intégrité : *la dépendance fonctionnelle.*

2.7 Redondances internes

Notion de dépendance fonctionnelle

LIVRE					
NUMERO	TITRE	AUTEUR	ISBN	DATE_ACHAT	EMPL
1029	L'humanité perdue	Finkelkraut A.	2 02 033300 7	14/10/2008	F3
1030	L'humanité perdue	Finkelkraut A.	2 02 033300 7	14/10/2008	F3
1032	Mercure	Nothomb A.	2 253 14911 X	14/10/2008	G5
1045	Eva Luna	Allende I.	2 253 05354 6	22/2/2009	F3
1067	Mercure	Nothomb A.	2 253 14911 X	24/2/2009	G5
1022	Mercure	Nothomb A.	2 253 14911 X	3/10/2008	G6

ISBN \rightarrow TITRE, AUTEUR

**si deux lignes ont la même valeur de ISBN,
alors elles ont aussi les mêmes valeurs de TITRE et AUTEUR**

On dit que :

- il existe une dépendance fonctionnelle de ISBN vers TITRE et AUTEUR
- ISBN *détermine* ou *est un déterminant* de TITRE et AUTEUR
- TITRE et AUTEUR *dépendent de*, ou *sont déterminés par* ISBN

2.7 Redondances internes

Notion de dépendance fonctionnelle

Deux observations

1. un identifiant détermine toutes les colonnes de la table
2. si un groupe de colonne détermine chaque colonne de la table, il constitue un identifiant de la table

NUMERO —→ **TITRE, AUTEUR, ISBN, DATE_ACHAT, EMPL**

2.7 Redondances internes

Comment détecter les situations de redondance ?

Réponse

Il y a redondance interne dès qu'il existe un déterminant qui n'est pas un identifiant de la table

Une dépendance fonctionnelle dont le déterminant n'est pas un identifiant est dite **anormale**

ISBN est un déterminant dans LIVRES mais n'en est pas un identifiant. Il entraîne donc des redondances internes.

2.7 Redondances internes

Comment corriger les situations de redondance ?

Réponse

En décomposant la table T en deux fragments :


T1(déterminant, déterminé)

T2(déterminant, résidu)

T2.déterminant **clé étrangère** vers T1

OUVRAGE		
ISBN	TITRE	AUTEUR
2 02 033300 7	L'humanité perdue	Finkelkraut A.
2 253 14911 X	Mercure	Nothomb A.
2 253 05354 6	Eva Luna	Allende I.

EXEMPLAIRE			
NUMERO	ISBN	DATE_ACHAT	EMPL
1029	2 02 033300 7	14/10/2008	F3
1030	2 02 033300 7	14/10/2008	F3
1032	2 253 14911 X	14/10/2008	G5
1045	2 253 05354 6	22/2/2009	F3
1067	2 253 14911 X	24/2/2009	G5
1022	2 253 14911 X	3/10/2008	G6



2.7 Redondances internes

Dernières remarques

1. Une table qui est le siège d'une dépendance fonctionnelle anormale est dite **non normalisée**
2. Une table sans dépendance fonctionnelle anormale est dite **normalisée**
3. Décomposer une table de manière à éliminer ses dépendances anormales consiste à **normaliser cette table**
4. Il est essentiel que toutes les tables d'une base de données soient normalisées
5. Il est possible qu'une table qui est le siège de dépendances fonctionnelles anormales ne comporte pas de redondance à certains moments. Il ne s'agirait que d'un accident statistique!