

Le modèle relationnel

1. Le modèle de données
2. Les contraintes d'intégrités
3. Les langages
 - algèbre, calcul
 - SQL

Le modèle de données

- **simple**

schéma = description des colonnes d'une table

instance = lignes de la table

- **formel**

théorie des ensembles, fonctions ...

logique

1970 : Codd

Un petit exemple avant les définitions

FILM	Titre	Metteur-en-scène	Acteur
	Speed 2	Jan de Bont	S. Bullock
	Speed 2	Jan de Bont	J. Patric
	Speed 2	Jan de Bont	W. Dafoe
	Marion	M. Poirier	C. Tetard
	Marion	M. Poirier	M-F Pisier

schéma de base de données Cinéma

FILM(Titre, Metteur-en-scène, Acteur)

CINE(Nom-Ciné, Adresse, Téléphone)

PROGRAMME(Nom-Ciné, Titre, Horaire)

Schéma de relation

- un nom de relation **R** **FILM**
 - un ensemble **Att(R)** (fini) d'attributs **A**, ...
Titre, **Metteur-en-scène**, **Acteur**
- pour chaque attribut
- un domaine (ens. de valeurs) **Dom(A)**
Dom(Titre) = chaînes de caractères de longueur 20

Notation :

R(A,B,C) **FILM(Titre, Metteur-en-scène, Acteur)**

arité de **R** = cardinalité de **Att(R)**

schéma de base de données relationnelle

un ensemble (fini) de schémas de relation (i.e. plusieurs descriptions de table).

Instance d'un schéma de relation $\mathbf{R}(A_1, A_2, \dots, A_n)$

Qu'est-il autorisé de mettre dans une ligne d'une table ?

- une ligne de la table = un **n-uplet** sur (A_1, A_2, \dots, A_n) .
 (a_1, a_2, \dots, a_n) telles que
pour tout $i = 1..n$, $a_i \in \text{Dom}(A_i)$
n-uplet = fonction, élt du produit cartésien

Notation : $u[A_i] = a_i =$ composante de u sur A_i .

Qu'est-il autorisé de mettre dans une table ?

- une **instance** \mathbf{r} de $\mathbf{R} =$
un ensemble **fini** de n-uplets sur (A_1, \dots, A_n)
un sous ensemble **fini** de $\text{Dom}(A_1) \times \dots \times \text{Dom}(A_n)$.

Instance d'un schéma de relation – exercice

- instance = un des contenus possibles de la table !!!!!!!!!

Considérons le schéma de relation STOCK(Habit,Taille) où
 $\text{Dom}(\text{Habit}) = \{\text{Tailleur}, \text{Costume}\}$ et $\text{Dom}(\text{Taille}) = \{38, 40\}$.

Combien d'instances de cardinalité 0, 1, 2, 3, 4, 5, ...

Instance d'une base de données $\{R_1, \dots, R_m\}$

Qu'est-il autorisé de mettre dans une base de données ?

- m instances de schéma de relation r_1, \dots, r_m (une pour chaque R_i) telles que r_i est une instance de R_i .

Instance de Base de Données – exemple

FILM	Titre	Metteur-en-scène	Acteur
	Speed 2	Jan de Bont	S. Bullock
	Speed 2	Jan de Bont	J. Patric
	Speed 2	Jan de Bont	W. Dafoe
	Marion	M. Poirier	C. Tetard
	Marion	M. Poirier	M-F Pisier

CINE	Nom-Ciné	Adresse	Téléphone
	Français	9, rue Montesquieu	05 56 44 11 87
	Gaumont	9, c. G-Clémenceau	05 56 52 03 54
	Trianon	6, r. Franklin	05 56 44 35 17
	UGC Ariel	20, r. Judaique	05 56 44 31 17

Instance de Base de Données – exemple

PROGR.	Nom-Ciné	Titre	Horaire
	Français	Speed 2	18h00
	Français	Speed 2	20h00
	Français	Speed 2	22h00
	Français	Marion	16h00
	Trainon	Marion	18h00

Contraintes d'intégrité – Dépendances

- base \leftrightarrow informations
- information \neq données des tables
- information \approx données \models **contraintes**

Différents types de contraintes :

contraintes **structurelles** (liées au modèle)

composante d'un n-uplet : $u[A_i] \in \text{Dom}(A_i)$

contraintes **liées à l'application**

contraintes **statiques**

chaque instance doit satisfaire ces propriétés

contraintes **dynamiques**

évolution de la base de données

Contraintes : exemple 1

une seule adresse et un seul numéro de téléphone par cinéma.

CINE	Nom-Ciné	Adresse	Téléphone
	Français	9,rue Montesquieu	05 56 44 11 87
	Français	9,rue Montesquieu	08 01 68 04 45
	UGC	20, rue Judaique	05 56 44 31 17
	UGC	20, rue Judaique	08 01 68 70 14
	UGC	22, rue Judaique	08 01 68 70 14

L'instance est **correcte** (définition d'une instance de **CINE**).

L'instance est **incohérente** (ne satisfait pas la propriété)

Contraintes : exemple 2

Plusieurs adresses et plusieurs lignes de téléphone par cinéma.

CINE	Nom-Ciné	Adresse	Téléphone
	Français	9,rue Montesquieu	05 56 44 11 87
	Français	9,rue Montesquieu	08 01 68 04 45
	UGC	20, rue Judaique	05 56 44 31 17
	UGC	20, rue Judaique	08 01 68 70 14
	UGC	22, rue Judaique	08 01 68 70 14

L'instance est **correcte** (définition d'une instance de **CINE**).

L'instance est **incohérente** (ne satisfait pas la propriété)

Contraintes : exemple 2

Plusieurs adresses et **plusieurs** lignes de téléphone par cinéma.

CINE	Nom-Ciné	Adresse	Téléphone
	Français	9,rue Montesquieu	05 56 44 11 87
	Français	9,rue Montesquieu	08 01 68 04 45
	UGC	20, rue Judaique	05 56 44 31 17
	UGC	20, rue Judaique	08 01 68 70 14
	UGC	22, rue Judaique	05 56 44 31 17
	UGC	22, rue Judaique	08 01 68 70 14

L'instance est **correcte** (définition d'une instance de **CINE**).

L'instance est **cohérente** (satisfait la propriété)

Contraintes : exemple 3

Les films **projetés** sont des films **répertoriés**.

PROGRAMME	Nom-Ciné	Titre	Horaire
	Français	Speed 2	18h00
	Français	Speed 2	20h00
	Français	Western	18h00
	Français	Western	20h00

FILM	Titre	Metteur-en-scène	Acteur
	Speed 2	Jan de Bont	S. Bullock
	Speed 2	Jan de Bont	J. Patric
	Speed 2	Jan de Bont	W. Dafoe
	Marion	M. Poirier	C. Tetard
	Marion	M. Poirier	M-F Pisier

L'instance est **correcte** (définition d'une instance de bd).

L'instance est **incohérente** (ne satisfait pas la propriété)

Dépendances Fonctionnelles(exemple 1)

- contraintes importantes : notion de clé
- mono-relation schéma : $\mathbf{R}(A_1, \dots, A_n)$
- syntaxe :

dépendance fonctionnelle : $\mathbf{X} \longrightarrow \mathbf{Y}$

dire : X détermine Y

$\mathbf{X}, \mathbf{Y} \subseteq \{A_1, \dots, A_n\}$.

- sémantique :

$\mathbf{r} \models \mathbf{X} \longrightarrow \mathbf{Y}$

dire : l'instance \mathbf{r} satisfait $\mathbf{X} \longrightarrow \mathbf{Y}$

pour tout couple de n-uplets u et v de \mathbf{r} on a :

$(\forall \mathbf{A} \in \mathbf{X} \ u[\mathbf{A}] = v[\mathbf{A}]) \Rightarrow (\forall \mathbf{A} \in \mathbf{Y} \ u[\mathbf{A}] = v[\mathbf{A}])$

Exemple 1 : (reprise)

- **une seule** adresse par cinéma.

Nom-Ciné \longrightarrow Adresse

vérification des dépendances fonctionnelles :

$u = (\text{UGC}; 20, \text{ rue Judaique}; 05\ 56\ 44\ 31\ 17)$

$v = (\text{UGC}; 22, \text{ rue Judaique}; 08\ 01\ 68\ 70\ 14)$

$u[\text{Nom-Ciné}] = v[\text{Nom-Ciné}] = \text{UGC}$ **mais**

$u[\text{adresse}] = 20 \text{ rue Judaique}$ et $v[\text{adresse}] = 22 \text{ rue Judaique}$

- **un seul** numéro de téléphone par cinéma.

Nom-Ciné \longrightarrow Téléphone

vérification de la dépendance fonctionnelle :

les mêmes n-uplets

Exemple Exercice

La relation (instance) ci-dessous satisfait-elle les dfs suivantes ?

- $A \longrightarrow C$
- $C \longrightarrow A$
- $AB \longrightarrow D$
- $A \longrightarrow A$

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_2	c_1	d_2
a_2	b_2	c_2	d_2
a_2	b_3	c_2	d_3
a_3	b_3	c_2	d_4

- dfs triviales (satisfaites pas toute instance r):
 $X \longrightarrow Y$ avec $Y \subseteq X$.

Schéma avec dépendances fonctionnelles

$(\mathbf{R}, \mathcal{F})$: \mathbf{R} est un schéma de relation et
 \mathcal{F} un ensemble de dfs

Une instance d'un schéma avec dfs $(\mathbf{R}, \mathcal{F})$ est une instance de \mathbf{R} qui satisfait chaque df de \mathcal{F} .

Remarque : – **Important** –

Lorsqu'on considère un schéma **avec** dfs, **seules** les instances satisfaisant les dfs sont considérées. En fait donner un ens. de dfs sert à éliminer les instances de \mathbf{R} qui ne correspondent pas à une situation “réelle”.

Implication de dfs ... un exemple ...

FILM(Titre, M-e-S, Acteur) avec (df) Titre \longrightarrow M-e-S

- $\forall r$ instance de **FILM** avec (df),
 $r \models$ Titre Acteur \longrightarrow M-e-S.

PROGbis(Titre, Résumé, Salle, Nom-Ciné) avec les 2 dfs :

(df1) Salle Nom-Ciné \longrightarrow Titre (df2) Titre \longrightarrow Résumé

Un ciné projette un seul film dans une salle

Un seul résumé est proposé pour chaque film

- $\forall r$ instance de **PROGbis** avec df1 et df2,
 $r \models$ Salle Nom-Ciné \longrightarrow Résumé

Implication de dépendances

Les instances d'un schéma avec dfs $(\mathbf{R}, \mathcal{F})$ satisfont bien sûr les dfs de \mathcal{F} (par définition) mais pas seulement.

• Conception

Le client énonce quelques propriétés

Le concepteur découvre que propriétés $\Rightarrow P$

\hookrightarrow Le client valide / ne valide pas P

\hookrightarrow Le concepteur utilise P pour optimiser / corriger

\hookrightarrow Clés

La Question (difficile):

Quelles sont toutes les dfs satisfaites par les instances de $(\mathbf{R}, \mathcal{F})$?

i.e. Quelles sont les dfs impliquées par \mathcal{F} ?

Implication de dépendancesschéma $(\mathbf{R}, \mathcal{F})$

- Y est la **fermeture de X** dans \mathcal{F} , noté X^+

Pour un X donné :

le plus grand ensemble Y tel que \mathcal{F} implique $X \longrightarrow Y$

- Un **algorithme simple** pour calculer la **fermeture de X** :

entrée : $(\mathbf{R}, \mathcal{F})$ et X

sortie : X^+

initialisation : $X^+ := X$

itération : si $Z \subseteq X^+$ et $(Z \longrightarrow A) \in \mathcal{F}$

alors $X^+ := X \cup \{ A \}$

Fermeture de X : des exemples

CINE(Nom-Ciné, Adresse, Téléphone) avec les dfs
Nom-ciné \longrightarrow Adresse et Nom-ciné \longrightarrow Téléphone.

$$(\text{Nom-Ciné})^+ = \{ \text{Nom-Ciné, Adresse, Téléphone} \}$$

R(A, B, C, D) avec les dfs A \longrightarrow B et BC \longrightarrow D

$$A^+ = AB$$

$$B^+ = B \quad C^+ = C$$

$$AC^+ = ABCD$$

AC est une clé

Clé d'un schéma avec dfs (\mathbf{R} , \mathcal{F})

- **Super-clé** : X tel que $X^+ = \text{Att}(\mathbf{R})$
 - **Clé** : une super-clé X telle qu'
aucun sous-ens. de X n'est une super-clé
-

CINE(Nom-Ciné, Adresse, Téléphone) avec Nom-ciné \longrightarrow Adresse
et Nom-ciné \longrightarrow Téléphone admet une clé : Nom-Ciné.

CINE(Nom-Ciné, Adresse, Téléphone) **sans dfs** admet une clé
(triviale): Nom-Ciné, Adresse, Téléphone.

Dépendances multivaluées (exemple 2)

- contraintes importantes : notion d'indépendance
- mono-relation : $R(A_1, \dots, A_n)$
- syntaxe

$$X \twoheadrightarrow Y \mid Z$$

dire : X multidétermine Y et Z

$$X, Y \text{ et } Z \subseteq \text{Att}(R) \text{ et } Z = \text{Att}(R) - XY$$

- sémantique

$$\mathbf{r} \models X \twoheadrightarrow Y \mid Z$$

si $\forall u, v \in \mathbf{r}$ tels que $u[X]=v[X]$, il existe

$$t \in \mathbf{r} \text{ tel que } t[XY]=u[XY] \text{ et } t[Z]=v[Z]$$

$$w \in \mathbf{r} \text{ tel que } w[XY]=v[XY] \text{ et } w[Z]=u[Z]$$

Dépendances multivaluées ... par exemple ...

	X	Y	Z
u	$a_1 \dots a_i$	$c_{i+1} \dots c_j$	$c_{j+1} \dots c_n$
v	$a_1 \dots a_i$	$d_{i+1} \dots d_j$	$d_{j+1} \dots d_n$
t	$a_1 \dots a_i$	$c_{i+1} \dots c_j$	$d_{j+1} \dots d_n$
w	$a_1 \dots a_i$	$d_{i+1} \dots d_j$	$a_{j+1} \dots a_n$

Exemple 2

- Nom-Ciné $\longrightarrow \longrightarrow$ Adresse | Téléphone

Un exemple de DMV

Nom \longrightarrow Cours | Sport

ETUDIANT	Nom	Cours	Sport
	Jean	BD	Natation
	Jean	BD	Football
	Jean	BD	Parapente
	Jean	Algo	Natation
	Jean	Algo	Football
	Jean	Algo	Parapente
	Jean	TeleInf	Natation
	Jean	TeleInf	Football
	Jean	TeleInf	Parapente
	Jean	C++	Natation
	Jean	C++	Football
	Jean	C ++	Parapente
	Jean	Gestion	Natation
	Jean	Gestion	Football
	Jean	Gestion	Parapente

REDONDANCE

Dépendances multivaluées : un exemple (suite)

- Suppression des redondances – Nouveau schéma

ET_COURS	Nom	Cours
	Jean	BD
	Jean	Algo
	Jean	TeleInf
	Jean	C++
	Jean	Gestion

ET_SPORT	Nom	Sport
	Jean	Natation
	Jean	Football
	Jean	Parapente

Dépendances d'inclusion (exemple 3)

- contraintes importantes : ISA, clé étrangère
- multi-relation 2 schémas : R_1, R_2
- syntaxe

$$R_1[A_1 .. A_n] \subseteq R_2[B_1 .. B_n]$$

$$A_1 .. A_n \subseteq \text{Att}(R_1) \text{ et } B_1 .. B_n \subseteq \text{Att}(R_2)$$

- sémantique

$$r_1, r_2 \models R_1[A_1 .. A_n] \subseteq R_2[B_1 .. B_n]$$

$\forall u \in r_1$, il existe $v \in r_2$ tel que

$$u[A_k] = v[B_k] \text{ pour } k=1..n.$$

il faut que les domaines de A_k et B_k soient compatibles

CINE[Nom-Cine] \subseteq CINE[Téléphone] n'a pas de sens !!!

Dépendances d'inclusion Exemples

- Exemple 3 : **PROGRAMME**[Titre] \subseteq **FILM**[Titre]

- Clé étrangère / contrainte de référence

R(... ABC ...) et **S**(... A'B'C' ...)

A est compatible avec A', B est ...

ABC est une clé de **R**

S[A', B', C'] \subseteq **R**[A, B, C]

déclaration avec SQL2

Conception de schéma

Anomalies de mise à jour.

Titre	M-e-S	Acteur	N-Ciné	Adr.	Tél	Heure
Speed	J. de B.	Bullock	Français	9, M.	..1187	18h
Speed	J. de B.	Bullock	Français	9, M.	..1187	20h
Speed	J. de B.	Patric	Français	9, M.	..1187	18h
Speed	J. de B.	Patric	Français	9, M.	..1187	20h
Speed	J. de B.	Bullock	Français	9, M.	..0445	18h
Speed	J. de B.	Bullock	Français	9, M.	..0445	20h
Speed	J. de B.	Patric	Français	9, M.	..0445	18h
Speed	J. de B.	Patric	Français	9, M.	..0445	20h
Speed	J. de B.	Bullock	Trianon	6, F.	..3117	16h
Speed	J. de B.	Patric	Trianon	6, F.	..3117	16h
:	:	:	:	:	:	:

Anomalies de mise à jour

- Insertion

n-uplet incomplet

complété par des valeurs spéciales (nulles)

ne pas stocker l'info

- Suppression

perte d'information

- Modification

mise à jour coûteuse

- éliminer les redondances

- propriétés des schémas de relations (**forme normale**)

- restructuration du schéma d'une base de données

3 ème Forme Normale

- $(\mathbf{R}, \mathcal{F})$ est sous **3ème forme normale** ssi pour toute df $X \longrightarrow A$ non triviale impliquée par \mathcal{F} une des deux conditions ci-dessous est vérifiée :

X est une super-clé

A appartient à une clé

Forme Normale de Boyce Codd

- $(\mathbf{R}, \mathcal{F})$ est sous **forme normale de Boyce-Codd** ssi pour toute df $X \longrightarrow A$ non triviale impliquée par \mathcal{F} la condition ci-dessous est vérifiée :

X est une super-clé

Relation non sous 3FN et redondance

Conf-Press(**Nom-Ciné**, **Salle**, **Horaire**, **Titre**) avec les dfs :

Nom-Ciné, **Titre** \longrightarrow **Salle**

Les conférences de presse pour un film ont lieu dans une salle.

Nom-Ciné, **Salle** \longrightarrow **Horaire**

Une seule conférence de presse par jour dans une salle.

Conf-Press	Nom-Ciné	Salle	Horaire	Titre
	Trianon	-1-	20h	Western
	Trianon	-1-	20h	Marion
	Trianon	-2-	16h	Speed2

Répétition de l'heure de la conférence pour le film Trianon

Une meilleure modélisation

Conf-Presse-Film(Nom-Ciné, Salle, Titre) avec la df

Nom-Ciné, Titre \longrightarrow Salle

Conf-Presse-Horaire(Nom-Ciné, Salle, Horaire) avec la df

Nom-Ciné, Salle \longrightarrow Horaire

Conf-Presse-Film	Nom-Ciné	Salle	Titre
	Trianon	-1-	Western
	Trianon	-1-	Marion
	Trianon	-2-	Speed2

Conf-Presse-Horaire	Nom-Ciné	Salle	Horaire
	Trianon	-1-	20h
	Trianon	-2-	16h

Forme normale de Boyce-Codd – un contre-exemple

FILM-débutant(Titre, M-e-S, Acteur) avec les dfs :

Titre, Acteur \longrightarrow M-e-S M-e-S, Acteur \longrightarrow Titre

Titre \longrightarrow M-e-S M-e-S \longrightarrow Titre

FILM-débutant	Titre	M-e-S	Acteur
	Speed2	J. de Bont	S. Bullock
	Speed2	J. de Bont	J. Patric
	Speed2	J. de Bont	W. Dafoe

Les colonnes Titre et M-e-S de cette table sont redondantes !

Une meilleure modélisation

FILM-deb-M-e-S(Titre, M-e-S) avec les dfs

Titre \longrightarrow M-e-S

M-e-S \longrightarrow Titre

FILM-deb-Acteur(Titre, Acteur) (sans df)

FILM-deb-M-e-S	Titre	M-e-S
	Speed2	J. de Bont

FILM-deb-Acteur	Titre	Acteur
	Speed2	S. Bullock
	Speed2	J. Patric
	Speed2	W. Dafoe

4ème forme normale dfs + mvds

- pas d'informations indépendantes dans une même relation

$(\mathbf{R}, \mathcal{F})$ est sous 4ème forme normale ssi

il est sous forme normale de Boyce-Codd et

pour toute dmvs $X \twoheadrightarrow Y$ impliquée par \mathcal{F} et non triviale,

X est une clé.

ETUDIANT(Nom, Cours, Sport) avec les dmvs

Nom \twoheadrightarrow Cours

Nom \twoheadrightarrow Sport

clé = Nom, Cours, Sport

pas sous 4ème forme normale

- Une meilleure modélisation

ET_COURS(Nom, Cours) **ET_SPORT**(Nom, Sport)

Décomposition (normalisation) / Dénormalisation

- **Décomposition** : suppression des redondances - schéma sous FN
optimisation des mises à jour

$$\begin{array}{ccc} R(A,B,..) & \implies & R_1(A_1,B_1,..) \quad \dots \quad R_k(A_k,B_k,..) \\ \mathcal{F} & & \mathcal{F}_1 \quad \dots \quad \mathcal{F}_k \end{array}$$

- **Dénormalisation** : regroupement - schéma non sous 3ème FN
optimisation de requêtes fréquentes (jointures)

$$\begin{array}{ccc} R_1(A_1,B_1,..) & R_2(A_2,B_2,..) & \implies R(A,B,..) \\ \mathcal{F}_1 & \mathcal{F}_2 & \mathcal{F} \end{array}$$

Décomposition – par l'exemple

BD-Cine(Titre, M-e-S, Acteur, Nom-Ciné, Adr, Tél) avec les dfs

Titre \longrightarrow M-e-S

Nom-Ciné \longrightarrow Tél Adresse

Titre \longrightarrow Nom-Ciné

Décomposition

{Titre, M-e-S, Acteur}

{Nom-Ciné, Titre}

{Nom-Ciné, Adresse, Tél}

↓

↓

↓

R₁(Titre, M-e-S, Acteur)

R₂(Nom-Ciné, Titre)

R₃(Nom-Ciné, Adresse, Tél)

Titre \longrightarrow M-e-S

Titre \longrightarrow Nom-Ciné

Nom-Ciné \longrightarrow Tél Adresse

Une instance de BD-Cine

BD-Cine	Titre	M-e-S	Acteur	Nom-Ciné	Adr.	Tél
	Speed 2	J. de B.	Bullock	Français	9, M.	..1187
	Speed 2	J. de B.	Patric	Français	9, M.	..1187
	⋮	⋮	⋮	⋮	⋮	⋮
	Marion	Poirier	Tetard	UGC	20, J.	..3117
	Marion	Poirier	Pisier	UGC	20, J.	..3117
	⋮	⋮	⋮	⋮	⋮	⋮

La décomposition de r sur R_1, R_2, R_3

r_1		
Titre	M-e-S	Acteur
Speed 2	J. de B.	Bullock
Speed 2	J. de B.	Patric
⋮	⋮	⋮
Marion	Poirier	Tetard
Marion	Poirier	Pisier
⋮	⋮	⋮

r_2	
Nom-Ciné	Titre
Speed 2	Français
⋮	⋮
Marion	UGC
⋮	⋮

r_3		
Nom-Ciné	Adr.	Tél
Français	9, M.	..1187
UGC	20, J.	..3117
⋮	⋮	⋮

Décomposition et perte d'information

$R(\text{Titre, Nom-Ciné, Tél})$ (sans dépendance)

Instance r		
Titre	Nom-Ciné	Tél
Speed 2	Français	..1187
Speed 2	Français	..1188
Marion	Français	..1187

Décomposition sur (Titre, Nom-Ciné) et (Nom-Ciné, Tél)

r_1		r_2	
Titre	Nom-Ciné	Nom-Ciné	Tél
Speed 2	Français	Français	..1187
Marion	Français	Français	..1188

Perte d'information

Décomposition de r

r_1

r_2

Titre	Nom-Ciné	Nom-Ciné	Tél
Speed 2	Français	Français	..1187
Marion	Français	Français	..1188

Recomposition par Jointure

Titre	Nom-Ciné	Tél
Speed 2	Français	..1187
Speed 2	Français	..1188
Marion	Français	..1187
Marion	Français	..1188

Décomposition sans perte d'information et dfs

Soit $(\mathbf{R}(XYZ), \mathcal{F})$ un schéma avec dfs.

La décomposition sur XY et XZ est sans perte d'information si

l'une condition ci-dessous est vraie :

$X \longrightarrow Y$ est impliquée par \mathcal{F}

$X \longrightarrow Z$ est impliquée par \mathcal{F}

$\mathbf{R}(A,B,C,D)$ avec la df $AB \longrightarrow C$.

(Exemple–Exercice)

Décomposition sur ABC et ABD .

\mathbf{r} sur \mathbf{R}

A	B	C	D
a ₁	b ₁	c ₁	d ₁
a ₁	b ₁	c ₁	d ₂
a ₁	b ₂	c ₁	d ₁
a ₁	b ₂	c ₁	d ₃
a ₂	b ₃	c ₂	d ₄

\mathbf{r}_1 sur \mathbf{R}_1

A	B	C
a ₁	b ₁	c ₁
a ₁	b ₂	c ₁
a ₂	b ₃	c ₂

\mathbf{r}_2 sur \mathbf{R}_1

A	B	D
a ₁	b ₁	d ₁
a ₁	b ₁	d ₂
a ₁	b ₂	d ₁
a ₁	b ₂	d ₃
a ₂	b ₃	d ₄

Décomposition et perte de dépendance

R(Titre Acteur Nom-Ciné) avec les dfs

Nom-Ciné \longrightarrow Titre Titre Acteur \longrightarrow Nom-Ciné

Décomposition :

R₁(Nom-Ciné ,Titre) **R**₂(Nom-Ciné, Acteur)

Nom-Ciné \longrightarrow Titre

Instance r satisfaisant les dfs

	Titre	Acteur	Nom-Ciné
	Speed2	Bullock	Français
	Speed2	Patric	Trianon

Décomposition de r

Nom-Ciné	Titre	Nom-Ciné	Acteur
Français	Speed2	Français	Bullock
Trianon	Speed2	Trianon	Patric

Insertion de (Trianon, Dafoe) dans R_2

Nom-Ciné	Titre	Nom-Ciné	Acteur
Français	Speed2	Français	Bullock
Trianon	Speed2	Trianon	Patric
		Trainon	Dafoe

Recomposition par Jointure

Titre	Acteur	Nom-Ciné
Speed2	Bullock	Français
Speed2	Patric	Trianon
Speed2	Dafoe	Trianon

Algorithme de Décomposition SPI en FNBC

Entrée : (R, \mathcal{F})

Sortie : schéma de base de données BD en FNBC

Etape 1. $BD := \{ (R, \mathcal{F}) \}$

Etape 2. Répéter

choisir dans BD un schéma $(S(V), \mathcal{G})$ non en FNBC

choisir une df $X \longrightarrow Y$ telle que X n'est pas clé

soit Z tel que $V=XYZ$

remplacer $(S(V), \mathcal{G})$ par

$(S_1(XY), \mathcal{G}_{XY})$ et $(S_2(XZ), \mathcal{G}_{XZ})$

jusqu'à ce que tous les schémas de BD en FNBC

Attention : Calcul de toutes les dfs impliquées par \mathcal{F}

Décomposition SPI en FNBC

Conf-Press(Titre Acteur Nom-Ciné) avec les dfs

Nom-Ciné \longrightarrow Titre Titre Acteur \longrightarrow Nom-Ciné

- Schéma non sous FNBC
- Seule Décomposition Possible:

\mathbf{R}_1 (Nom-Ciné ,Titre) \mathbf{R}_2 (Nom-Ciné, Acteur)

Nom-Ciné \longrightarrow Titre

- Décomposition SPI
- Décomposition non SPD

choisir entre :

FNBC avec **incohérence**

SPI et SPD avec **redondance**

Contraintes - Formes Normales - Décomposition

Dépendances fonctionnelles, multivaluées, de jointure, d'inclusion ...

1FN	valeurs atomiques des attributs
2FN	1FN + tout attribut non clé dépend
3FN	2FN + pas de dfs entre attributs non clé
FNBC	partie gauche de df super-clé
4FN	FNBC + partie gauche de mvd = super-clé

Décomposition SPI, SPD

Compromis Mise à Jour – Requêtes