

Sujet de TD n°1 BASES DE DONNÉES

Correction

Modèle relationnel, décomposition, pertes de données et de dépendances

Attention : vous ferez la question 5 comme exercice personnel une fois que l'on aura assez avancé dans le cours.

EXERCICE 1

Soit un schéma relationnel constitué d'une seule relation :

R (Id-Cours, Id-Etudiant, Age, Note)

et des deux dépendances fonctionnelles suivantes :

Id-Cours, Id-Etudiant → Note Id_Etudiant → Age.
--

QUESTIONS

1. Donner quelques exemples de tuples correspondant à la relation R.

Id-cours	Id-etudiant	Age	Note
C1	E1	18	15
C1	E2	20	16
C2	E1	18	12
C3	E1	18	12

2. Indiquer les clés candidates de la relation R

Une clé candidate c'est un attribut ou groupe d'attributs minimal qui peut identifier de façon unique chaque tuple de la relation

La valeur d'une clé candidate est distincte pour chaque tuple :

- Id-cours, Id-Etudiant, Age et Note pris séparément ne sont pas des clés candidates
- [Id-cours, Id-Etudiant, Age, Note] identifie de façon unique chacun des tuples mais n'est pas minimal
- Il semble que [Id-cours, Id-Etudiant] est clé candidate

Une clé candidate c'est un attribut ou ensemble d'attributs minimal en DF avec les autres attributs de la relation :

- [Id-cours, Id-Etudiant] est en DF élémentaire avec *Note* (1^{ère} DF de l'hypothèse)
- [Id-cours, Id-Etudiant] est en DF (mais pas élémentaire) avec *Age* (2^{ème} DF de l'hypothèse)

[Id-cours, Id-Etudiant] est donc une clé candidate.

3. Citer les anomalies et les redondances qui se trouvent dans la relation R

Comme on le voit dans les tuples qui ont été utilisés dans l'exemple, l'âge est répété pour chaque cours mettant en jeu le même étudiant... C'était prévisible compte tenu que [Id-cours, Id-Etudiant] est en DF (mais pas élémentaire) avec Age

4. Décomposer la relation R afin de supprimer les anomalies.

Qu'est-ce qu'il faut décomposer ? Les relations qui ont des dépendances autres que les dépendances fonctionnelles élémentaires entre la clé primaire et les autres attributs

Comment ?

Soit (si la DF à décomposer est élémentaire)

- En créant une relation dont la clé est la partie gauche de la DF et les autres attributs de cette relation, sa partie droite.
- En supprimant dans la relation initiale les attributs en partie droite de la DF.

Soit (si la DF à décomposer est non élémentaire)¹ :

- En créant une relation dont la clé est le « bout de la partie gauche » en DF élémentaire avec les attributs de la partie droite de la DF et les autres attributs de la relation, sa partie droite.
- En supprimant la partie droite de la DF dans la relation initiale.

Ici : **R (Id-Cours, Id-Etudiant, Age, Note)** devient donc :

- **R1 (Id-Cours, Id-Etudiant, Note)**
- **R2 (Id-Etudiant, Age)**

On a :

- supprimé dans la relation initiale (qui est devenue R1), la DF entre [Id-cours, Id-Etudiant] et Age
- créé la relation R2 dont la clé candidate est [Id-Etudiant] et qui a comme attribut Age.

5. Vérifier que la décomposition est sans perte de données (le vérifier expérimentalement en faisant une jointure puis en le démontrant à l'aide du théorème de HEATH) et sans perte de dépendances

Si on suit les définitions Il faut que la décomposition se fasse :

- Sans perte de données : la jointure naturelle des relations issues de la décomposition d'une relation R doit être une relation dont le contenu est équivalent à celui de la relation R
- Sans perte de dépendance : on peut alors retrouver (notamment par transitivité) toutes les DF de R à partir des relations issues de la décomposition

¹ On a vu en cours que pour une dépendance non élémentaire si on commence par s'occuper de la dépendance élémentaire qui rend justement la dépendance non élémentaire alors on se retrouve toujours à décomposer par rapport à des dépendances élémentaires.

Vérification de non perte de dépendance :

- On retrouve dans R2 la dépendance **Id_Etudiant** → **Age**
- On retrouve dans R1 la dépendance **Id-Cours, Id-Etudiant** → **Note**

Vérification de non perte de donnée :

On peut faire une première vérification en réalisant la jointure naturelle suivante :

R1 J{Id-Etudiant} R2

Id-cours	Id-etudiant	Note
C1	E1	15
C1	E2	16
C2	E1	12
C3	E1	12

Id-etudiant	Age
E1	18
E2	20

Cette définition permet de tout de suite dire qu'il y a perte de donnée si le contenu de R et de la relation issue de la jointure sont différent : « Soit une relation **R(A,B,C)** où A, B et C sont des ensembles d'attributs disjoints, avec **B** → **C** alors **R(A,B,C) = R[A,B] J{B} R[B,C]** »

Sinon il faut appliquer le théorème de HEATH : « Toute relation R(X, Y, Z) est décomposable sans perte d'information en

- R1 = projection (X, Y) de R et
- R2 = projection (X, Z) de R

S'il y a dans R une dépendance fonctionnelle de X vers Y ($X \rightarrow Y$) »

EXERCICE 2

Soit la relation :

Commande (No- Commande, No-Produit, Quantité-Commandée, No-Client, No-Représentant)
--

et les dépendances fonctionnelles suivantes :

No-Commande, No-Produit → Quantité-Commandée, No-Client, No-Représentant No-Commande → No-Client, No-Représentant No-Client → No-Représentant
--

QUESTIONS

1. Donner quelques exemples de tuples correspondant à la relation R.

No-Commande	No-Produit	Quantité-Commandée	No-client	No-Représentant
CO1	PR1	10	CL1	RE1
CO1	PR2	15	CL1	RE1
CO2	PR2	10	CL3	RE2
CO2	PR3	11	CL3	RE2

CO3	PR3	11	CL4	RE1
-----	-----	----	-----	-----

2. Indiquer les clés candidates de la relation R

- No-Commande, No-Produit, Quantité-Commandée, No-Client, No-Représentant pris séparément ne sont pas des clés candidates
- [No-Commande, No-Produit, Quantité-Commandée, No-Client, No-Représentant] identifie de façon unique chacun des tuples mais n'est pas minimal
- [No-Commande, No-Produit] semble être clé candidate

Une clé candidate c'est un attribut ou ensemble d'attributs minimal en DF avec les autres attributs de la relation :

- [No-Commande, No-Produit] est en DF (élémentaire) avec Quantité-Commandée, No-Client, No-Représentant (1^{ère} DF)
- Remarque : [No-Commande, No-Produit] est aussi en DF (non élémentaire) avec No-Client, No-Représentant
[No-Commande, No-Produit] est donc clé candidate

3. Citer les anomalies et les redondances qui se trouvent dans la relation R

Les DFs suivantes induisent des anomalies (les parties gauches ne correspondent pas à la clé candidate) :

- DF1 : No-Commande \rightarrow No-Client, No-Représentant
Duplication du numéro de client et du numéro de représentant chaque fois qu'un même produit est concerné par une commande.
- DF2 : No-Client \rightarrow No-Représentant
Duplication du numéro de représentant pour chaque commande concernant le même client.

4. Décomposer la relation R afin de supprimer les anomalies.

Si on applique ce qui a été dit à l'exercice n°1 :

On cherche d'abord à enlever DF1 :

- R1 (No-Commande, No-Produit, Quantité-Commandée) *correspond à une commande*
R1 contient seulement une DF qui a pour partie gauche la totalité de la clé.
- R2 (No-Commande, No-Client, No-Représentant)
 - R2 contient toujours une DF qui n'a pas pour partie gauche la totalité de la clé ; il faut donc décomposer R2 car R2 introduit des anomalies : duplication du numéro de représentant pour chaque commande concernant le même client.
 - R2' (No-Commande, No-Client) *correspond au fait qu'une commande est réalisée par un client et un seul.*
R2' contient seulement une DF qui a pour partie gauche la totalité de la clé.
 - R2'' (No-Client, No-Représentant) *correspond à l'affectation d'un représentant à un client (un client a un seul représentant qui s'occupe de lui)*
R2'' contient seulement une DF qui a pour partie gauche la totalité de la clé.

On cherche d'abord à enlever DF2 :

- R1 (No-Commande, No-Produit, Quantité-Commandée, No-Client)

- R1 contient toujours une DF qui n'a pas pour partie gauche la totalité de la clé ; il faut donc décomposer R1 car R1 introduit des anomalies : duplication du numéro de client pour chaque produit de la même commande.
 - R1' (No-Commande, No-Produit, Quantité-Commandée)
 - R1' (No-Commande, No-Client)
- R2 (No-Client, No-Représentant)

En partant de DF1 ou de DF2 on aboutie à la même décomposition (ce n'est pas toujours le cas) mais en passant par des relations intermédiaires différentes.

5. Vérifier que la décomposition est sans perte de données (le vérifier expérimentalement en faisant une jointure puis en le démontrant à l'aide du théorème de HEATH) et sans perte de dépendances

On montrera comme dans l'exercice n°1 qu'il n'y a pas de perte de règles de dépendances. Il reste à vérifier qu'il n'y a pas de perte de données en faisant les jointures naturelles

EXERCICE 3

Soit la relation :

R (Fournisseur, Adresse, Raison-Sociale, no-Produit, Libellé-Produit, Quantité, Prix, No-Commande, Délai, Date)

et les dépendances fonctionnelles suivantes :

No-Commande → Fournisseur, Délai, Date
Fournisseur → Raison-Sociale, Adresse
No-Commande, no-Produit → Quantité
No-Produit, Fournisseur → Prix
No-Produit → Libellé-Produit

QUESTIONS

1. Donner quelques exemples de tuples correspondant à la relation R.

Fournisseur	Adresse	Raison sociale	No-Produit	Libellé-Produit	Quantité	Prix	No-Commande	Délai	Date
F1	A1	R1	PO1	LIBP1	20	105	CO1	15	D1
<u>F1'</u>	<u>A1'</u>	<u>R1'</u>	PO2	<u>LIBP2'</u>	30	110	CO1	<u>15'</u>	<u>D1'</u>
F3	A3	R3	PO2	LIBP2	25	101	CO2	10	D2
<u>F3'</u>	<u>A3'</u>	<u>R3'</u>	PO3	LIBP3	20	120	CO2	<u>10'</u>	<u>D2'</u>
F1	<u>A1''</u>	<u>R1''</u>	PO1	<u>LIBP1'</u>	40	<u>105'</u>	C03	<u>15'</u>	D3

2. Indiquer les clés candidates de la relation R

- Aucun attribut n'est à lui seul clé candidate
- Bien sur l'ensemble des attributs identifie de façon unique chacun des tuples mais n'est pas minimal
- [No-Commande, No-Produit] est en DF élémentaire avec *Quantité* (1^{ère} DF de l'hypothèse)

- [No-Commande, No-Produit] est en DF (mais pas élémentaire) avec *les autres attributs*
 - De plus : No-Commande \rightarrow Fournisseur, Délai, Date
 - De plus : Fournisseur \rightarrow Raison Sociale, Adresse (par transitivité : No-Commande \rightarrow Raison Sociale, Adresse)Et donc No-Commande \rightarrow Fournisseur, Délai, Date, Raison Sociale, Adresse
Et donc [No-Commande, No-Produit] \rightarrow Fournisseur, Délai, Date, Raison-Sociale, Adresse, Quantité (mais pas élémentaire)
 - De plus No-Produit \rightarrow Libellé-ProduitEt donc [No-Commande, No-Produit] \rightarrow Libellé-Produit (pas élémentaire)
Et donc [No-Commande, No-Produit] \rightarrow Fournisseur, Délai, Date, Raison-Sociale, Adresse, Libellé-Produit, Quantité (mais pas élémentaire)
 - Si on applique l'axiome d'Armstrong de pseudo-transitivité :
Si $X \rightarrow Y$ et $WY \rightarrow Z$ alors $WX \rightarrow Z$
 $X = \text{No-Commande}$, $Y = \text{Fournisseur}$, $W = \text{No-Produit}$, $Z = \text{prix}$
On sait que :
No-Commande \rightarrow Fournisseur et No-Produit, Fournisseur \rightarrow Prix,
Donc : No-Commande, No-Produit \rightarrow PrixEt donc [No-Commande, No-Produit] \rightarrow Fournisseur, Délai, Date, Raison-Sociale, Adresse, Libellé-Produit, Prix, Quantité (mais pas élémentaire)
- [No-Commande, No-Produit] est une clé candidate

3. Citer les anomalies et les redondances qui se trouvent dans la relation R

On note de nombreuses anomalies dans le tableau qui résultent des dépendances fonctionnelles suivantes :

DF1 : No-Commande \rightarrow Fournisseur, Délai, Date

DF2 : Fournisseur \rightarrow Raison-Sociale, Adresse

DF3 : No-Produit, Fournisseur \rightarrow Prix

DF4 : No-Produit \rightarrow Libellé-Produit

On constate donc que :

- Le fournisseur, le délai et la date sont dupliquée chaque fois qu'une commande contient plusieurs produits
- La raison sociale et l'adresse sont dupliquées chaque fois qu'apparaît le même fournisseur dans une commande
- Le prix est dupliqué si dans plusieurs commandes apparaît le même produit et le même fournisseur
- Chaque fois qu'un produit apparaît dans une commande son libellé est dupliqué

4. Décomposer la relation R afin de supprimer les anomalies.

R (Fournisseur, Adresse, Raison-Sociale, no-Produit, Libellé-Produit, Quantité, Prix, No-Commande, Délai, Date) se décompose en :

- R1(Adresse, Raison-Sociale, no-Produit, Libellé-Produit, Quantité, Prix, No-Commande)
- R2(No-Commande, Fournisseur, Délai, Date) – DF1

Le problème est que l'on a perdu une DF (DF2). Il faut donc essayer de faire autrement !

R (Fournisseur, Adresse, Raison-Sociale, no-Produit, Libellé-Produit, Quantité, Prix, No-Commande, Délai, Date) se décompose en :

R1 (Fournisseur, no-Produit, Libellé-Produit, Quantité, Prix, No-Commande, Délai, Date)

R2 (Fournisseur, Raison-Sociale, Adresse) – DF2

R1 se décompose en :

R1' (no-Produit, Libellé-Produit, Quantité, Prix, No-Commande)

R1'' (No-Commande, Fournisseur, Délai, Date) – DF1

Le problème est que l'on a perdu une DF (DF3).

Il semble que si on procède dans l'ordre DF4, DF3, DF2, DF1 on n'ait pas de perte de dépendance car les parties droites n'influencent pas les DF (partie gauche) suivantes (à vérifier)

5. Vérifier que la décomposition est sans perte de données (le vérifier expérimentalement en faisant une jointure puis en le démontrant à l'aide du théorème de HEATH) et sans perte de dépendances

Voir 4). Il reste à vérifier qu'il n'y a pas de perte de données en faisant les jointures naturelles

EXERCICE 4

Soit la relation :

R (A, B, C, D, E)

avec les dépendances fonctionnelles suivantes :

A → B, C, D, E

C, D → E

E, C → B

QUESTIONS

1. Donner quelques exemples de tuples correspondant à la relation R.

A	B	C	D	E
A1	B1	C1	D1	E1
A2	B2	C1	D2	E2
A3	B2	C1	D3	E2
A4	B4	C2	D1	E3

2. Indiquer les clés candidates de la relation R
C'est [A] qui est clé candidate (il n'y a pas plus petite)
Il n'y a qu'elle car A n'apparaît en partie droite d'aucune DF
3. Citer les anomalies et les redondances qui se trouvent dans la relation R

On note des anomalies dans le tableau qui résultent des dépendances fonctionnelles suivantes :

- C, D → E et

- $E, C \rightarrow B$

Et donc :

- Chaque tuple ayant le même valeur de C et D implique la duplication de E
- Chaque tuple ayant le même valeur de E et C implique la duplication de B

4. Décomposer la relation R afin de supprimer les anomalies.

R (A, B, C, D, E) est décomposé en :

- R1 (A, C, D)
- R2 (C, D, E, B) qui se décompose en R2'(C, D, E) et R2''(E, C, B)

5. Vérifier que la décomposition est sans perte de données (le vérifier expérimentalement en faisant une jointure puis en le démontrant à l'aide du théorème de HEATH) et sans perte de dépendances

A partir de R1 (A, C, D), R2'(C, D, E) et R2''(E, C, B)

On retrouve immédiatement :

$A \rightarrow C, D$ et $C, D \rightarrow E$ et $E, C \rightarrow B$

Il reste à retrouver $A \rightarrow B, E$

Si on applique l'axiome de transitivité d'Armstrong : Si $A \rightarrow C, D$ et $C, D \rightarrow E$ alors $A \rightarrow E$

Et Donc $A \rightarrow C, D, E$ (il ne reste plus que $A \rightarrow B$)

Si on applique l'axiome d'augmentation d'Armstrong : Si $X \rightarrow Y$ alors $XZ \rightarrow YZ$

Si $E, C \rightarrow B$ alors $E, C, D \rightarrow D, B$

Si $A \rightarrow C, D, E$ et $E, C, D \rightarrow D, B$ alors (par transitivité) : $A \rightarrow D, B$

Donc $A \rightarrow C, D, E, B$

Donc aucune perte de dépendance

Il reste à vérifier qu'il n'y a pas de perte de données en faisant les jointures naturelles