Relations d'Armstrong, redondances et perte d'information

UE fondements des bases de données - TD5

1. **Les relations d'Armstrong**

   Soit un schéma de relation R et un ensemble F de DF sur ce schéma. Une *relation d'Armstrong* pour F est une relation r sur R vérifiant *exactement* F^+, ni plus ni moins. Son intérêt est de fournir un exemple représentatif au concepteur, pour l'aider à normaliser ou détecter des anomalies.

   Nous allons voir sur un exemple comment construire une telle relation. Soit $R = ABCDE$ et $F = \{A \to B, C; D \to E; C \to D\}$.

   (a) Calculez l'ensemble des fermés de F, défini par $Cl(F) = \{X^+, X \subseteq R\}$. Pour celà, utilisez une exploration systématique des sous-ensembles de R en commençant par ceux de taille 1, puis 2, etc... Pensez à ne pas visiter les sur-ensembles des clés (cf TD précédent)

   On trouve les fermés suivants : $ABCDE, BCDE, B, CDE, DE, E, BDE,$

   (b) Le principe est alors, pour chaque $X \in Cl(F) - R$, de s'assurer que deux tuples au moins prennent les mêmes valeurs sur X et des valeurs différentes ailleurs.

   Appliquez l'algorithme suivant :

Algorithm 1 Relation d'Armstrong pour un ensemble de DF

Require: Un schéma de relation R, un ensemble de DF F sur R.
Ensure: Une relation d'Armstrong r pour F sur R.

for all $A \in R$ **do**
  $t[A] = 0$;
end for
$r = \{t\}$;
$i = 1$;
for all $X \in Cl(F) - R$ **do**
  for all $A \in R$ **do**
    if $A \in X$ **then** $t[A] = 0$
    else $t[A] = i$;
  end for
  $r = r \cup \{t\}$;
  $i := i + 1$;
end for
Retourner r ;

On trouvera, en les considérant dans l'ordre ou je les ai donnés :

A	B	C	D	E
0	0	0	0	0
1	0	0	0	0
2	0	2	2	2
3	3	0	0	0
4	4	4	0	0
5	5	5	5	0
6	0	6	0	0

2. On souhaite créer une base de données de recettes de cuisine, décrites comme suit :

   Une recette est identifiée par son numéro. Elle a un nom et un type particulier (soupe, entrée, dessert, ...). Elle utilise un ou plusieurs ingrédients (carottes, viande de boeuf, poivre, ...). Une ingrédient est identifié par son numéro et à un nom. Pour chaque ingrédient dans une recette, on précisera sa quantité et l'unité de mesure pour l'exprimer.

(a) Donner le schéma de la *relation universelle* permettant de modéliser ces données, c'est à dire la liste des attributs de l'application regroupés dans une seule relation R. Dressez l'inventaire des DF valides sur ce schéma.

$R = (NumR, NomR, TypeR, NumI, NomI, Qte)$

$F = \{NumR \rightarrow NomR, TypeR; NumI \rightarrow NomI; NumR, NumI \rightarrow Qte\}$

(b) Construisez une BD d'Armstrong pour F.

(c) Exhibez les problèmes de redondance à partir de cet exemple. La DF $NumR, NumI \rightarrow qte$ engendre-t-elle un problème de redondance ?

Ils savent qu'un problème de redondance a lieu lorsqu'une DF prends deux fois la même valeur à gauche et à droite. la dernière DF ne pose pas de problème, puisque la partie gauche est une clé.

(d) Proposez de façon intuitive une décomposition sans pertes et sans redondance de R.

$R_1 = (\overline{NumR}, NomR, TypeR)$, $R_2 = (\overline{NumR, NumI}, Qte)$, $R_3 = (\overline{NumI}, NomI)$.

(e) On veut maintenant ajouter à cette modélisation qu'à chaque recette correspond un ensemble d'ustensiles. Cette information se traduit-elle par une nouvelle DF ?

Non

(f) Que pensez-vous alors de la décomposition suivante :

$R_1 = (\overline{NumR}, NomR, TypeR)$, $R_2 = (\overline{NumR, NumI}, Qte)$, $R_3 = (\overline{NumI}, NomI)$, $R_4 = (\overline{NumR, Ustensile})$.

Les laisser réfléchir - Cette décomposition engendre une perte de jointure... Il suffit pour ca d'exhiber la relation suivante sur R :

No. Tuple	NumR	NomR	TypeR	NumI	NomI	Qte	Ustensile
1	1	Couscous	Plat principal	1	Pois Chiches	0,5kg	Couscoussière
2	1	Couscous	Plat principal	2	Agneau	1kg	Plat

On peut vérifier quette relation ne viole aucune DF, elle est donc valide.

On construit alors les relations r_1, r_4 en prenant les projections correspondantes. Ensuite, on fait la jointure naturelle entre ces relations. On obtient :

No. Tuple	NumR	NomR	TypeR	NumI	NomI	Qte	Ustensile
1	1	Couscous	Plat principal	1	Pois Chiches	0,5kg	Couscoussière
1	1	Couscous	Plat principal	1	Agneau	1kg	Couscoussière
1	1	Couscous	Plat principal	1	Pois Chiches	0,5kg	Plat
2	1	Couscous	Plat principal	2	Agneau	1kg	Plat

Cette relation est différente de la première, il y a donc perte de jointure et la décomposition n'est pas correcte.