



Responsables : Z. Kedad, B. Finance

Tous documents (papier) autorisés, Exercices indépendants, Durée : 2h

EXERCICE 1

Soit le schéma cible suivant :

Client (ID, Nom, Adresse)

Commande (ID_commande, Date, *Code_produit*, *ID*, Nom_produit, Quantité)

Produit (Code_produit, Nom_produit, Prix unitaire, Nom_fournisseur)

Et soient les trois sources de données suivantes :

Source S1 :

Client_IDF (ID, Nom, Adresse), Client_BDR (ID, Nom, Adresse)

Commande (ID_Cde, Date, *Code_produit*, *ID*, Quantité)

(Les deux tables clients représentent des clients de régions différentes)

Source S2 :

Produit (Code_Produit, Nom_produit, Prix unitaire)

Source S3 :

Fournisseur (Code_Fournisseur, nom)

Achat (Code_produit, *Code_Fournisseur*)

Les clés primaires sont soulignées, les clés étrangères sont en italique.

Questions :

1. On souhaite définir les mappings permettant de construire un système d'intégration et ce selon une approche GAV. Pour chaque objet du schéma cible :
 - a. Donner le graphe d'opérations en précisant la règle d'intégration permettant de déduire ces opérations (jointures).
 - b. Donner ensuite le ou les chemins de calcul possibles.
 - c. S'il existe plusieurs chemins de calcul, donner une façon possible de traiter ces chemins pour cet exemple, en justifiant.
2. On suppose que l'utilisateur pose la requête suivante sur le schéma cible :

Select Client.Nom, Commande.Nom_produit, Produit.Nom_fournisseur

From Client, Commande, Produit

Where client.ID = Commande.ID, Commande.Code_produit = Produit.Code_produit

Expliquer les étapes de la réécriture de cette requête en considérant les mappings obtenus en question 1.

3. Toujours en considérant les mappings définis en question 1, on souhaite maintenir le système d'intégration en état de fonctionnement même si des évolutions surviennent au niveau des sources :
 - a. Quel est l'impact de la suppression de la table Produit dans la source 2 ?
 - b. Quel est l'impact de la suppression de la table Client_IDF dans la source 1 ?
 - c. Quel est l'impact de la suppression de la table Client dans le schéma cible ?
 - d. Donner les principales étapes pour prendre en compte la suppression d'une table dans une source de données et adapter les mappings.
4. On suppose maintenant que l'on souhaite définir les mappings selon une approche LAV.
 - a. Donner l'ensemble des mappings selon cette approche.
 - b. Quel est l'impact de la suppression de la table Client dans le schéma cible ?
 - c. Quel est l'impact de la suppression de la table Client_IDF dans la source S1 ?
5. Dites quels sont les cas où l'approche LAV est plus adaptée, et dans lesquels il faut préférer l'approche GAV (justifier).

Remarques : notez qu'un graphe d'opérations peut être réduit à un seul sommet. Un chemin de calcul peut également être réduit à un seul sommet du graphe d'opération

EXERCICE 2

Une société de service informatique (SSII) intervient pour le compte d'entreprises clientes dans la gestion de leurs projets informatiques. Elle effectue des interventions auprès de ces entreprises. Une intervention consiste à affecter une ressource humaine (employé de la SSII) pour une certaine durée pour la réalisation d'un projet. Chaque intervention est chiffrée et correspond à un montant payé par l'entreprise cliente (le montant de chaque intervention représente le coût d'une journée de travail d'un employé par le nombre de jour travaillé par l'employé).

La SSII souhaite analyser son chiffre d'affaire par entreprise cliente, par personne employée, par zone géographique et par projet. Elle souhaite disposer d'indicateurs pour identifier les entreprises avec lesquelles elle travaille le plus, ou encore les périodes auxquelles elle travaille le plus.

1. Proposer un schéma conceptuel pour un entrepôt de données permettant de supporter ces différentes analyses.
2. Si les données sont stockées sous forme de structure multidimensionnelle, donner la séquence d'opérateurs multidimensionnels permettant de :
 - a. retrouver le montant des interventions pour le projet P1.
 - b. retrouver le nombre de jours d'intervention par mois dans une région donnée.
 - c. retrouver le montant total facturé pour le projet P1 et la personne EMP1.

EXERCICE 4

L'objectif de cet exercice est de proposer un système de recommandation de livres pour les fêtes de Noël. Soient les 3 prédicats suivants, correspondants à des relations stockées dans une base de données relationnelles:

- $\text{Book}(x,y)$: indique que le livre x est du genre y . Ex : $\text{Book}(\text{'GamesOfTrones'}, \text{'fantastic'})$
- $\text{Friend}(x,y)$: indique que la personne x est ami de la personne y . Ex : $\text{Friend}(\text{'Tom'}, \text{'Jack'})$.
- $\text{Like}(x, y)$: indique que la personne x aime y .
 - o Ex : $\text{Like}(\text{'Tom'}, \text{'GameOfThrones'})$, $\text{Like}(\text{'Jack'}, \text{'fantastic'})$.

Ecrire en DATALOG les choses suivantes :

- 1- Les amis de mes amis sont mes amis.
- 2- Si un de mes amis aime un livre alors on me recommande de lire ce livre.
- 3- Si j'aime le genre 'x' ('fantastic') alors on me recommande de lire tous les livres de ce genre.

EXERCICE 5

Soit le programme **Datalog^{neg}** suivant :

$\text{Anc}(\text{'a'}, \text{'b'})$.

$\text{Anc}(\text{'b'}, \text{'c'})$.

$S(x, y) :- \text{Anc}(x, z), \text{Anc}(z, y)$.

$Q(x, z) :- S(x, y), \neg R(y), T(y, z)$.

$R(x) :- S(x, y), \neg Q(x, y)$.

1/ Peut-on calculer la sémantique de ce programme ? Si oui dites comment ? Donner les étapes et les explications nécessaires à son obtention.

EXERCICE 6

Exercice 2 (4p) : Soit le programme Datalog suivant pour aider au transport routier.

$\text{Road}(\text{'Paris'}, \text{'Lyon'})$.

$\text{Road}(\text{'Lyon'}, \text{'Marseille'})$.

$\text{Road}(\text{'Paris'}, \text{'Rouen'})$.

$\text{Road}(\text{'Rouen'}, \text{'Lyon'})$.

$\text{Link}(x, y) :- \text{Road}(x, y)$.

$\text{Link}(y, x) :- \text{Road}(x, y)$.

$\text{Link}(x, z) :- \text{Link}(x, y), \text{Link}(y, z), x \neq z$.

On cherche à répondre aux questions suivantes :

- Peut-on aller de Paris à Marseille par la route ?
- Peut-on aller du Havre à Marseille par la route ?

Q1 : Calculer le modèle (sémantique) du programme DATALOG permettant de répondre à ces 2 questions de façon formelle.