



Responsables : Z. Kedad, B. Finance

Tous documents (papier) autorisés, Exercices indépendants, Durée : 3h

---

## EXERCICE 1

Une entreprise souhaite analyser les ventes de ses produits, elle souhaite entre autres analyser les ventes par zone géographique pour ses différents types de produits. Elle souhaite également disposer d'indicateurs sur ses clients pour identifier les meilleurs. Un des besoins exprimés est l'identification des périodes de l'année où les produits sont le mieux vendus.

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 total des ventes du produit P1.
- b. retrouver le nombre d'unités vendues par jour dans une ville donnée tous produits confondus.
- c. retrouver le montant total des ventes pour chaque mois de l'année.

3/ On souhaite également analyser les achats de matières premières auprès de fournisseurs dans différentes villes. Modifier le schéma conceptuel précédent pour prendre en compte ce nouveau besoin.

## EXERCICE 2

Les systèmes d'intégration ont pour but de fournir une vue uniforme sur un ensemble de sources de données hétérogènes. Parmi les systèmes d'intégration existants, on trouve les systèmes de médiation et les entrepôts de données. La construction de ces systèmes suppose la résolution des conflits existant entre les sources de données, qui peuvent être hétérogènes, distribuées et autonomes.

1. Donner un exemple de conflit pouvant survenir au niveau du schéma de deux sources distinctes.
2. Donner un exemple de conflit pouvant survenir au niveau des instances de deux sources distinctes.
3. La conception d'un entrepôt de données suppose la sélection des agrégats à matérialiser. Quels sont les paramètres à prendre en compte pour faire ce choix.
4. Quelle est la différence entre le chargement initial d'un entrepôt de données et le rafraîchissement de l'entrepôt ?

### EXERCICE 3

Dans un système d'intégration, soient les relations cibles  $R(\underline{A}, B, C, D)$  et  $R'(\underline{B}, E)$  et deux sources de données S1 et S2.

S1 contient les tables :  $R1(\underline{A}, B)$ ,  $R2(\underline{A}, B)$ ,  $R3(\underline{C}, D)$ .

S2 contient les tables :  $R4(\underline{B}, E)$ ,  $R5(\underline{E}, D)$ ,  $R6(\underline{B}, D)$ .

(Les clés primaires sont soulignées)

Les mappings décrivent des expressions de calcul entre les relations sources et les relations cibles.

1. Donnez tous les mappings pour ce système en supposant une approche Global As View (GAV).
2. Donnez tous les mappings pour ce système en supposant une approche Local As View (LAV).
3. Montrer sur cet exemple qu'il est possible pour une relation d'être associée à deux mappings ayant des sémantiques différentes.
4. Dans le cas où une relation admet plusieurs mappings, donnez une façon possible de choisir le mapping qui sera utilisé au moment de la réécriture d'une requête posée par l'utilisateur au système d'intégration.
5. On suppose que le contenu des tables R5 et R6 évolue très fréquemment suite à des insertions ou des suppressions de tuples. mises à jour. Est-il préférable de choisir l'approche GAV ou l'approche LAV pour définir les mappings ? Pourquoi ?
6. On suppose maintenant que de nouvelles sources de données sont très fréquemment ajoutées au système. Est-il préférable de choisir l'approche GAV ou l'approche LAV pour définir les mappings ? Pourquoi ?
7. On suppose que les relations cibles sont fréquemment supprimées et de nouvelles relations cibles ajoutées. Est-il préférable de choisir l'approche GAV ou l'approche LAV pour définir les mappings ? Pourquoi ?

### EXERCICE 4

Mettre sous forme de clauses de Horn la formule logique suivante :

$(\exists X (\text{CHIEN}(X)) \wedge (\exists Y (\text{AIME}(X, Y) \vee \forall X \text{CONTENT}(X)))$

### EXERCICE 5

Soit le programme Datalog<sup>neg</sup> suivant :

OISEAU (Pegase).  
MARCHE (Pegase).  
OISEAU (Titi).  
PINGUOIN (x) <- OISEAU (x), MARCHE (x).

$\text{VOLE}(x) \leftarrow \text{OISEAU}(x), \neg \text{PINGUOIN}(x).$

Calculer la sémantique de ce programme. En donnant les étapes et les explications nécessaires à son obtention.

## EXERCICE 6

On considère les trois prédicats suivants :

$\text{Employe}(x).$

$\text{A-Comme-Responsable}(x,y).$

$\text{Responsable}(y).$

**Question 1:** Ecrire, en Datalog, la requête permettant de retourner tous les employés ayant le même responsable que “Smith”.

**Question 2:** Ecrire, en Datalog, la requête permettant de retourner toutes les personnes (les personnes regroupent les employés et les responsables)

**Question 3:** Ecrire, en Datalog, la requête permettant de retourner tous les responsables (direct ou indirect  $n+1$ ,  $n+2$ , etc...) pour chaque employé. On pourra au besoin définir une vue.

**Question 4:** Soient les deux vues suivantes :

$V1(x,y) :- \text{A-Comme-Responsable}(x,u), \text{A-Comme-Responsable}(u,y).$

$V2(x,y) :- \text{A-Comme-Responsable}(x,y), \text{Employe}(y).$

Dire en langage naturel ce que calcule chacune de ses vues.