

UNIVERSITÉ ASSANE SECK DE ZIGUINCHOR
UFR DES SCIENCES ET TECHNOLOGIES
DÉPARTEMENT D'INFORMATIQUE

CHAPITRE IV

OPTIMISATION DE REQUÊTES

ANNÉE ACADÉMIQUE : 2022 – 2023

FILIÈRE : INGÉNIERIE INFORMATIQUE

NIVEAU : LICENCE 3

SEMESTRE : 5

DR SERIGNE DIAGNE

PLAN DU COURS

Introduction

I. Arbre algébrique

1. Définition
2. Représentation des opérations de l'algèbre relationnelle

II. Optimisation de requêtes

1. Restructuration algébrique
2. Heuristiques d'optimisation

III. Exercice d'application

INTRODUCTION

- La plupart des SGBD relationnels modernes offrent des langages de manipulation basés sur SQL ;
- SQL est un langage non procédural qui utilise des opérateurs ensemblistes ;
- L'utilisateur définit les données qu'il veut visualiser sans fournir des algorithmes d'accès à ces données ;
- L'objectif de l'optimisation est de déterminer ces algorithmes d'accès appelés **plan d'exécution** ;
- Il est essentiel pour un système d'utiliser des plans d'exécution optimisés pour les requêtes les plus fréquentes ;
- Un plan d'exécution dépend du schéma interne de la base de données (particulièrement de l'existence d'index) et de la taille des tables.

INTRODUCTION

- Un optimiseur de requêtes transforme une requête exprimée dans un langage source (SQL par exemple) en un plan d'exécution ;
- Un plan d'exécution est une séquence d'opérations de bas niveau réalisant efficacement l'accès aux données ;
- 4 ➤ L'exécution d'une requête SQL par un SGBD suit les étapes suivantes :
 - ✓ **Analyse syntaxique** : Vérification de la syntaxique et traduction en opérations algébriques ;
 - ✓ **Contrôle de l'accès aux données** : Vérification des droits d'accès de l'utilisateur ;
 - ✓ **Optimisation** : Génération des plans d'exécution et choix du meilleur ;
 - ✓ **Exécution** : Compilation et exécution du plan choisi.
- Optimiser une requête c'est chercher le plan d'exécution optimal pour minimiser le coût de son exécution en ressources et diminuer la durée de son exécution ;
- Le système construit tous les arbres algébriques possibles pour la requête, puis évalue leur coût et enfin choisit celui qui a le plus petit coût d'exécution.

I. ARBRE ALGÈBRIQUE

I. 1. Définition

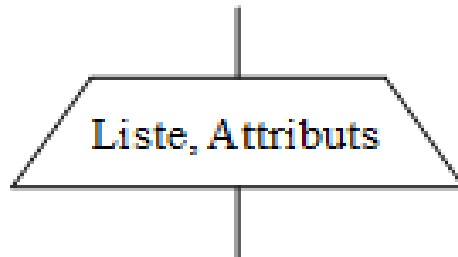
C'est la représentation graphique sous forme d'arbre d'une requête dans laquelle :

- les feuilles représentent les relations ;
- les nœuds intermédiaires représentent les opérateurs algébriques ;
- le nœud racine représente la relation résultat de la requête ;
- les arcs représentent les flux de données entre les opérations ;
- les tris sont représentés par des rectangles contenant les attributs sur lesquels ils portent ;
- les agrégats sont représentés par un rectangle contenant les attributs de la clause **Group By** et d'un autre rectangle contenant les attributs résultats calculés.

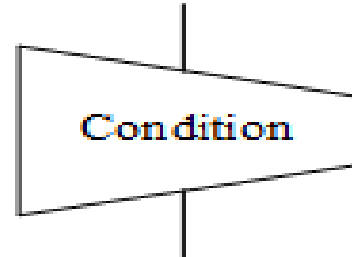
I. ARBRE ALGÉBRIQUE

I. 2. Représentation des opérations de l'algèbre relationnelle

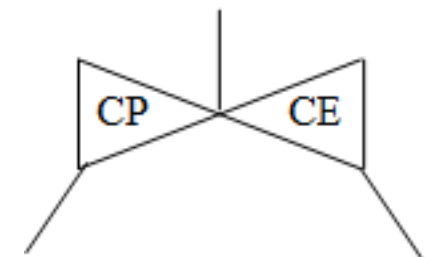
Projection



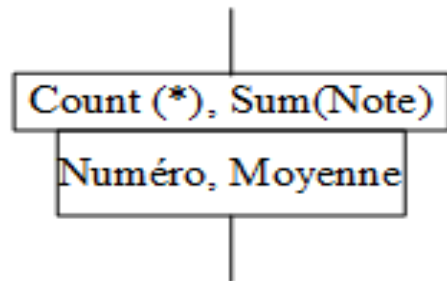
Sélection



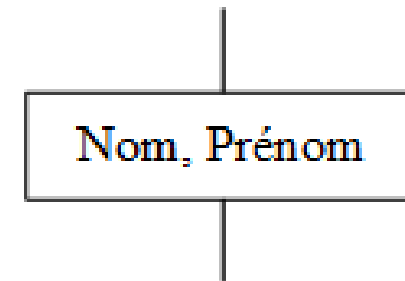
Jointure naturelle



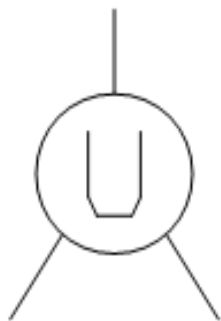
Tri



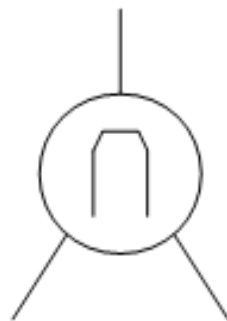
Agrégats



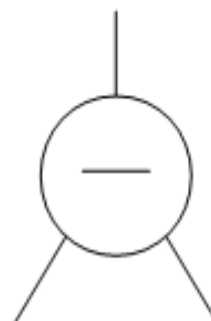
Union



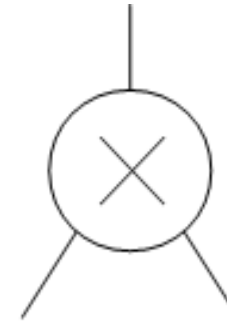
Intersection



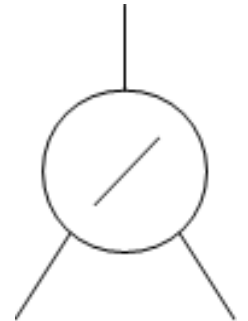
Différence



Produit cartésien



Division



II. OPTIMISATION D'ARBRES ALGÈBRIQUES

II. 1. Restructuration algébrique

Pour optimiser une requête on la réécrit pour obtenir une requête équivalente en utilisant :

- la commutativité et l'associativité de la jointure naturelle : elles permettent de changer l'ordre des jointures pour :
 - ✓ minimiser la taille des données à parcourir ;
 - ✓ Diminuer le nombre de comparaisons de valeurs à faire ;
- le regroupement des sélections : il permet d'effectuer plusieurs sélections en un seul parcours de la table au lieu d'une sélection par parcours ;
- la commutativité des sélections et jointures : elle permet d'effectuer les sélections avant les jointures ;
- la descente des projections en veillant à conserver les attributs qui seront affichées ou utilisés dans des conditions.

II. OPTIMISATION D'ARBRES ALGÈBRIQUES

II. 2. Heuristiques d'optimisation

L'ordonnancement des opérations se fait ensuite comme suit :

8

- on fait descendre les opérations réductrices (projection, sélection) le maximum possible vers les feuilles ;
- on retarde le plus possible les jointures et les produits cartésien ;
- on exécute d'abord à chaque fois que c'est possible les jointures qui font appel aux relations les moins volumineuses.

III. EXERCICE D'APPLICATION

Bus (Matricule, Marque, Version, Annee, Nb_Place, Consommation, Vitesse)

Chauffeur (Num_permis, Nom, Prenom, Age, Sexe, Adresse, Telephone, Annee_permis)

Passager (NIN, Nom, Prenom, Adresse, Telephone, Age, Sexe)

9 Voyage (Date, Heure_Depart, Ville_Depart, Heure_Arrive, Ville_Arrive, #Bus)

Effectuer (#Chauffeur, #Date, #Heure_Depart, #Ville_Depart)

Voyager (#Client, #Date, #Heure_Depart, #Ville_Depart)

I. Donner les arbres relationnels optimisés des requêtes suivantes :

1. Quelles sont les villes de départ et d'arrivée du passager Cheikh Gueye le 20/05/2020 ?
2. Quel est le nombre de places du bus conduit par Aminata Fall le 14/01/2021 à 10h ?
3. Quelle est la moyenne d'âge des passagers du bus conduit par Jean Ndong le 08/03/2021 à 18h ?
4. Matricules bus et Numéros de téléphone de chauffeurs ayant conduit un bus pris par Abdou GUEYE ?