

Tables

```
CREATE TABLE Etudiant(  
    NumEtud INT,  
    NomEtud VARCHAR(100),  
    PrenomEtud VARCHAR(100),  
    AdresseEtud VARCHAR(100),  
    Filiere VARCHAR(100),  
  
    PRIMARY KEY(NumEtud)  
)  
  
CREATE TABLE Option(  
    NumOption INT,  
    NomOption VARCHAR(100),  
    NomIntervenant VARCHAR(100),  
    VolHoraire REAL,  
    Credits INT,  
  
    n_ecole INT,  
  
    PRIMARY KEY(NumOption),  
)  
  
CREATE TABLE Inscription(  
    NumEtud INT,  
    NumOption INT,  
    AnneeI INT,  
    NoteObtenue REAL,  
  
    PRIMARY KEY(NumEtud, NumOption, AnneeI),  
    FOREIGN KEY(NumEtud) REFERENCES Etudiant,  
    FOREIGN KEY(NumOption) REFERENCES Option  
)
```

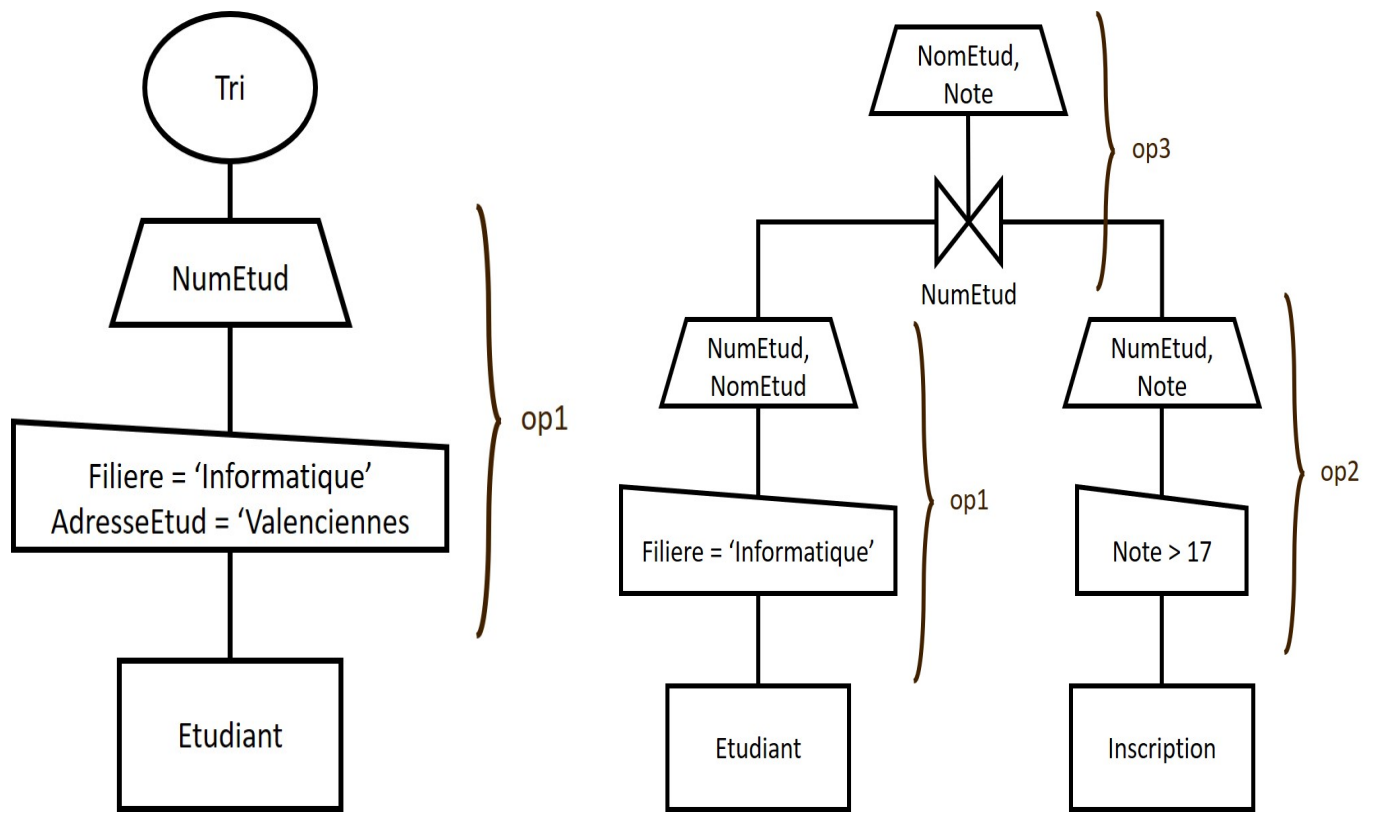
Requêtes

```
!R1  
Select E.NumEtud  
From Etudiant E  
Where  
    E.Filiere = 'Informatique'  
    And E.AdresseEtud = 'Valenciennes'  
Order by NumEtud ;  
  
!R2  
Select E. NomEtud, I.NoteObtenue  
From Etudiant E, Inscription I  
Where  
    I.NoteObtenue > 17  
    And E.Filiere = 'Informatique'  
    And E.NumEtud = I.NumEtud
```

Question 1 : Donner le ou les arbre(s) algébrique(s) optimisé(s) pour les requêtes R1 et R2

R1 :

R2 :



Question 2 : Donner pour chaque arbre les plans d'exécution possibles.

R1 : 2 plans d'exécution possibles :

- PE1 : RBS + TRI (inutile)
 - Ici le tri est inutile car la RBS ne change pas l'ordre et de base les tuples d'Etudiant sont triés par NumEtud
- PE2 : RI + TRI

R2 : 4 plans d'exécution possibles :

- PE1 : RBS + RBS + JBI
- PE2 : RI + RBS + JBI
- PE3 : RBS + RBS + JTF
- PE4 : RI + RBS + JTF

Question 3 : Calculer le coût des différents plans d'exécution en considérant que :

- la taille de tous les attributs numériques ou de type date est de 10 Ko
- la taille de tous les attributs de type chaîne de caractère est de 30 Ko
- la taille d'une page est de 1000 Ko
- la relation Etudiant possède 100000 tuples
- la relation Option possède 500 tuples
- la relation Inscription possède 600000 tuples
- la sélectivité ($I.NoteObtenue > 17$) est égale à 0,04
- 10% des étudiants habitent Valenciennes
- Il y a 25 filières représentées au sein de l'université, on suppose qu'elles sont équitablement réparties en nombre d'étudiants.
- Un index est créé sur l'attribut Filiere
- CPU = 0.1
- Buf = 10 (nombre de buffers consacrés aux algorithmes)

Relation 1 :

R1-PE1 :

- **op1 : RBS**

- $Coût = NT(S) \times CPU + NP(S)$
- $Coût = 100000 \times 0.1 + \frac{100000 \times 130}{1000}$
- $Coût = 23000$

Cout R1-PE1 : 23000

R1-PE2 :

- **op1 : RI**

- $Coût = Sélectivité(surl'index).NT(Rel).CPU + NP(Index)$
- $Coût = 0,04 \times 100000 \times 0,1 + \frac{34 \times 100000}{1000}$
- $Coût = 3800$

- **op2 : TRI**

- $NT_{Resop1} = 100000 \times 0,1 \times 0,04$
- $TT_{Resop1} = 10$
- $NP_{Resop1} = 4$
- $Coût_{Tri} = 2 \times 4 \log_4 \simeq 5$

Cout R1-PE2 : 3805

Relation 2 :

- **op1 :**

- $Coût_{op1}(RBS) = 23000$
- $Coût_{op1}(RI) = 3800$

- **op2 :**

- $Coût_{op2}(RBS) = 600000 \times 0,1 + \frac{600000 \times 40}{1000} = 84000$

- **op3 :**

- $NT_{Resop1} = 0,04 \times 100000 = 4000$ tuples
- $TT_{Resop1} = 30 + 10$ ko
- $NP_{Resop1} = \frac{40 \times 4000}{1000} = 160$ pages
- $NT_{Resop2} = 0,04 \times 600000 = 24000$ tuples
- $TT_{Resop2} = 10 + 10$ ko
- $NP_{Resop2} = \frac{20 \times 24000}{1000} = 480$ pages
- $Coût_{op3}(JBI) = (NT(Resop1) + NT(Resop2)) \times CPU + NP(Resop1) \times (1 + \frac{NP(Resop2)}{Buf}) = (4000 + 24000) \times 0,1 + 160 \times (1 + \frac{480}{10}) = 10640$
- $CoutTri(Resop1) = 0$
- $CoutTri(Resop2) = 2 \times NP(Resop2) \times \log_{Buf}(NP(Resop2)) = 2 \times 480 \times \log_{10}(NP(480)) \simeq 2574$
- $Coût_{op3}(JTF) = (NT(Resop1) + NT(Resop2)) \times CPU + CoutTri(S) + coutTri(Q) + NP(Resop1) + NP(Resop2) = (4000 + 24000) \times 0,1 + 0 + 257 + 160 + 480 = 6014$

Finalement :

- $Coût(PE1) = 23000 + 84000 + 10640 = 117640$
- $Coût(PE2) = 3800 + 84000 + 10640 = 98440$
- $Coût(PE3) = 23000 + 84000 + 6014 = 113014$
- $Coût(PE4) = 3800 + 84000 + 6014 + CoutTri(Resop1) = 94519$
 - On ajoute $CoutTri(Resop1)$ car après un RI, ce n'est pas trié. Donc on ne peut pas omettre $CoutTri(Resop1)$ dans $Coût_{op3}(JTF)$

Le plan d'exécution optimal est donc **PE4**.