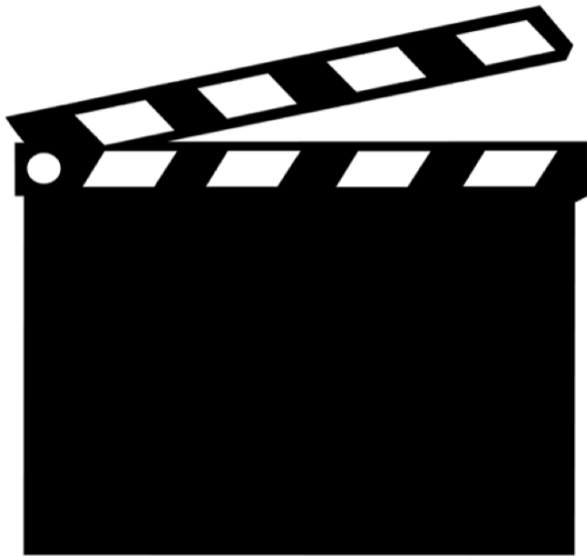


C018SA-W3-S2



Semaine 3 : Exécution et optimisation

1. Introduction
2. **Réécriture algébrique**
3. Opérateurs
4. Plans d'exécution
5. Tri et hachage
6. Algorithmes de jointure
7. Optimisation

De SQL vers une forme opératoire : l'algèbre

Titre des films parus en 1958, où l'un des acteurs joue le rôle de John Ferguson.

En SQL :

```
select titre
from   Film f, Role r
where  nom_role = 'Ferguson'
and    f.id = r.id_ilm
and    f.annee = 1958
```

Deux sélections (l'année, le rôle), une jointure, une projection.

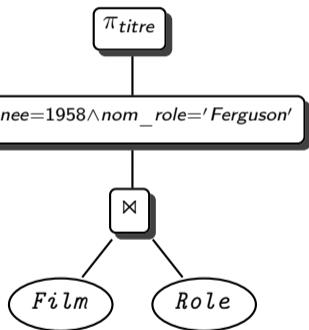
Le Plan d'Exécution Logique (PEL)

En algèbre, sous une forme normalisée projection-sélection-jointure :

$$\pi_{titre}(\sigma_{annee=1958 \wedge nom_role='Ferguson'}(Film \bowtie_{id=id_film} Role))$$

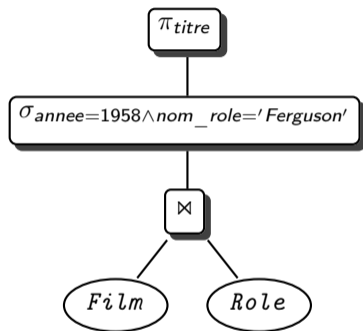
Le Plan d'Exécution Logique (PEL)

La même expression, sous forme d'un arbre.



Le Plan d'Exécution Logique (PEL)

La même expression, sous forme d'un arbre.



Est-ce le seul ? Non : des **réécritures** vont nous permettre d'en trouver d'autres.

Réécritures

Il y a plusieurs expressions **équivalentes** pour une même requête.

On peut explorer ces expressions grâce à des règles de réécriture. Exemple

1. **Commutativité des jointures** : $R \bowtie S \equiv S \bowtie R$
2. **Regroupement des sélections** : $\sigma_{A='a' \wedge B='b'}(R) \equiv \sigma_{A='a'}(\sigma_{B='b'}(R))$
3. **Commutativité de σ et de π** : $\pi_{A_1, A_2, \dots, A_p}(\sigma_{A_i='a'}(R)) \equiv \sigma_{A_i='a'}(\pi_{A_1, A_2, \dots, A_p}(R))$
4. **Commutativité de σ et de \bowtie** : $\sigma_{A='a'}(R[\dots A \dots] \bowtie S) \equiv \sigma_{A='a'}(R) \bowtie S$
5. etc.

L'application dirigée d'une règle d'équivalence (réécriture) transforme une expression e en une expression e' **équivalente**.

Ce que fait l'optimiseur

Trouve les expressions équivalentes, évalue leur coût et choisit la meilleure.

Important : on ne peut pas **énumérer** tous les plans possibles (trop long) : on applique des **heuristiques**.

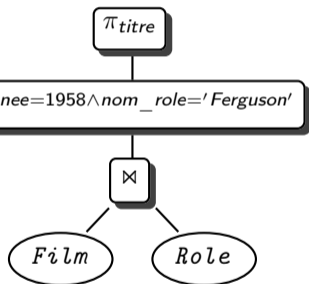
Heuristique classique : **réduire la taille des données**

- en filtrant les nuplets par des sélections
- en les simplifiant par des projections

dès que possible.

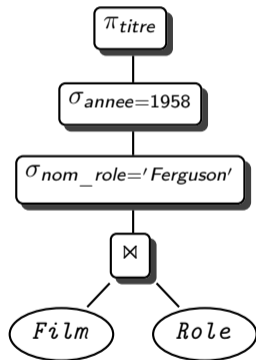
Illustration

Reprenons : le film paru en 1958 avec un rôle 'Ferguson'.



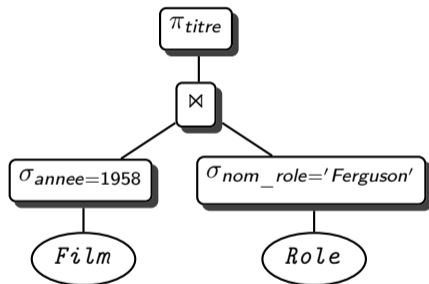
Illustration

Première étape : je sépare les sélections.



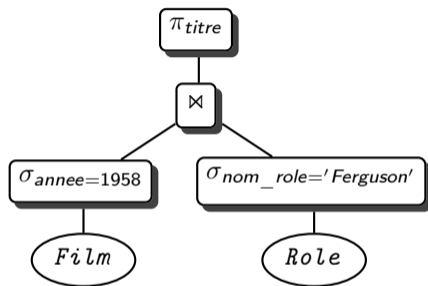
Illustration

Puis on pousse chaque sélection vers sa table.



Illustration

Puis on pousse chaque sélection vers sa table.



À ce stade : reste à choisir les chemins d'accès et les algorithmes de jointure.

Résumé : la réécriture algébrique

Principes essentiels :

1. L'algèbre permet d'obtenir une version opératoire de la requête.
2. Les équivalences algébriques permettent d'explorer un ensemble de plans.
3. L'optimiseur évalue le coût de chaque plan.

Bien retenir

- **Heuristique** : on ne peut pas **tout explorer**
- **Nécessaire mais pas suffisant** : il reste à choisir le bon algorithme pour chaque opération.

Résumé : la réécriture algébrique

Principes essentiels :

1. L'algèbre permet d'obtenir une version opératoire de la requête.
2. Les équivalences algébriques permettent d'explorer un ensemble de plans.
3. L'optimiseur évalue le coût de chaque plan.

Bien retenir

- **Heuristique** : on ne peut pas **tout explorer**
- **Nécessaire mais pas suffisant** : il reste à choisir le bon algorithme pour chaque opération.

Merci !