**UE SGBD 2019-2020**

**Projet Master**

**L'objectif du projet** est de comparer, dans des environnements variés, les coûts d'exécution de différents algorithmes de jointure.

Dans le cadre de ce projet et comme nous l'avons fait en cours, le coût d'exécution d'un algorithme sera exprimé uniquement par le **nombre de blocs** lus et/ou écrits pour l'exécution de cet algorithme.

Le calcul des coûts demandé n'est pas le résultat d'une évaluation théorique mais un calcul de coût réel d'exécution sur des tables existantes (que vous devrez créer) dans différentes situations (tables triées ou non, présence d'un ou plusieurs index...).

Pour faciliter votre tâche et avant d'écrire les algorithmes de jointure, vous devrez développer un **outil composé de briques logicielles** qui permettront l'écriture simplifiée de ces algorithmes en fournissant les opérations nécessaires à leur mise en œuvre (lecture/écriture de blocs, lectures/écriture de tuples, gestion de la mémoire cache..).

**Ce qui est à réaliser** :

- développer les **briques de base "a minima"** : ces briques se présenteront sous forme de classes d'objets dotés des seules propriétés et méthodes utiles à notre projet. Leur implémentation se fera au plus simple. Le développement des briques ne doit pas être "réaliste" mais juste permettre l'écriture des algorithmes des opérateurs de l'algèbre pour un calcul de leur coût. Par exemple l'implémentation d'un bloc pourra se faire sous la forme d'un texte de longueur fixe stocké dans un tableau de textes. Le numéro de bloc dans ce cas serait le numéro de ligne dans le tableau.

- **développer au moins 2 algorithmes de jointure** en utilisant les briques de base.

- **réaliser une interface** qui permette de créer des tables et des indexes puis de mettre en œuvre ces algorithmes et d'en calculer le coût d'exécution.

**Les briques de base** : classes "tables", "lignes"," indexes", "blocs", "mémoire cache", "buffer".

Les tables devront pouvoir être initialisées par une méthode de génération de tuples automatique.

**Les algorithmes de jointure** :

- produit cartésien, key-lookup, sort-merge, hashage, Omiecinski (jointure en présence de 2 indexes secondaires).

**Vous devrez rédiger un rapport** composé de :

- la présentation du projet

- la présentation des briques de base

- la présentation des algorithmes dont vous calculer le coût

- l’interprétation des résultats (expliquer les résultats obtenus avec les formules de coût)