



**INSTITUTO FEDERAL MINAS GERAIS (IFMG) - CAMPUS BAMBUÍ**  
**Banco de Dados II**  
**Prof. Marcos Roberto Ribeiro**

**Lista de Exercícios 9**

Beatriz Rodrigues de Oliveira Paiva

**Exercício 1:**

Considere uma seleção com apenas uma condição simples. Como tal operação é avaliada se a condição não envolve índices? E se a condição envolver índices, devemos sempre usá-los?

Será necessário analisar os tipos de índices, varrer todos os valores fora do índice. Podemos usá-la para criação de índices, eliminação de duplicadas e para alguns algoritmos de junção.

**Exercício 2:**

Explique os dois métodos de seleção sem disjunção. E no caso das seleções com disjunção, o que pode acontecer?

Primeira opção: Deve calcular o caminho mais seletivo, assim obtendo as tuplas e aplicar os demais componentes da condição para obter o resultado final.

Segunda opção: Obter os rids de todos os índices que satisfazem a condição, fazer uma interseção entre estes rids e aplicar o restante da condição de seleção. Caso seleções com disjunção, um único termo da disjunção pode provocar uma varredura de arquivo.

**Exercício 3:**

Explique as duas técnicas de avaliação de projeção (com eliminação de duplicatas) existentes. Qual das duas se sobressai? Podemos usar índices para avaliar tal operação?

Particionamento mas ordenação é mais usada, pode usar para tal operação sim.

Ordenação se sobressai, principalmente se tiver muitas duplicadas, pode sim usar índices, pois os valores possíveis já estão nos índices

**Exercício 4:**

É possível avaliar uma operação de junção usando uma equivalência com os operadores de produto cartesiano, seleção e projeção? isto é recomendável?

É possível, não é recomendável por conta do custo do produto cartesiano.

**Exercício 5:**

Explique como funcionam e compare o custo dos seguintes algoritmos de avaliação de junção:

(a) Junção de loops aninhados;

Por pegar página por página seu custo não é bom e para melhorar esse custo podemos ler uma página de cada tabela e fazer a junção das tuplas destas páginas. Seu custo  $M + M \times pR \times N$  E/S

(b) Junção de loops aninhados de bloco;

O algoritmo de junção de loops aninhados simples não utiliza o buffer de memória de forma eficiente. Supondo que hajam  $B$  páginas de memória livre. Podemos ler  $B/2$  páginas da relação externa, percorremos a relação interna usando uma página e usamos a última página para gravar o resultado. O custo será de  $M + N \times (M/B/2)$  E/S.

(c) Junção de loops aninhados indexados;

Só é possível utilizar essa junção caso esteja no índice, pois compara os índices das páginas. Custo  $M \times N E/S$ .

(d) Junção Sort-Merge;

As duas relações são ordenadas sobre o atributo de junção a ordenação agrupa as tuplas com o mesmo valor com o mesmo valor no atributo de junção em partições explorando este particionamento, comparamos as tuplas de R apenas com as tuplas de S que estão na mesma partição. Após a ordenação, começamos na primeira tupla de cada relação e avançamos quando o valor for menor do que da outra relação. Seu custo  $M \log M + N \log N + M \times N E/S$

(e) junção por hashing.

Possui duas fases:

Particionamento: Fazemos hash das duas relações no atributo de junção, usando a mesma função de hash  $h$ .

Correspondência: Comparamos as tuplas em uma partição de R somente com as tuplas na partição correspondente de S,

#### **Exercício 6:**

Descreva como as operações de conjunto podem ser avaliadas.

Em casos de interseção podemos utilizar a junção, cria o particionamento e faz as uniões nas partes.

#### **Exercício 7:**

Explique os métodos para avaliar as operações de agregação.

Caso exista o GROUPBY existem três estratégias possíveis: ordenação, hash e utilização dos índices.