



Instituto de Ciências Matemáticas e de Computação – USP
Programa de Ciências de Pós-graduação em Computação e Matemática Computacional

Lista de exercícios de SCC-5789 – Bases de Dados
Arquitetura de SGBDs Relacionais - Métodos de Acesso Físico
Segundo Semestre de 2012
Prof. Dr. Caetano Traina Júnior

Os exercícios desta lista se referem à seguinte definição da base de dados, a mesma base que tem sido usada durante as aulas com referência a um sistema de controle de matrículas de alunos, definida em Oracle:

Relação de alunos:

```
CREATE TABLE Alunos (  
  NUSP CHAR(10) PRIMARY KEY,  
  Nome CHAR(40),  
  NomeMae CHAR(40),  
  Idade DECIMAL(3),  
  Cidade CHAR(30),  
  Curso CHAR(30) )  
STORAGE INITIAL 20K NEXT 20K  
PCTFREE 10 PCTUSED 70;
```

Relação das matrículas efetuadas pelos alunos:

```
CREATE TABLE Matriculas (  
  Disciplina CHAR(8) NOT NULL,  
  NUSP CHAR(10) NOT NULL,  
  Frequencia DECIMAL(3),  
  PrimeiraAval DECIMAL(4,2),  
  SegundaAval DECIMAL(4,2),  
  Semestre CHAR,  
  Ano INT,  
  PRIMARY KEY (NUSP, Disciplina, Ano, Semestre) )  
STORAGE INITIAL 20K NEXT 20K  
PCTFREE 10 PCTUSED 70;
```

Assumimos que existam **80.000 alunos**, matriculados em **12.000 disciplinas**, cada aluno se matriculando numa **média de 8 disciplinas** por semestre. A Base usa páginas de 2K, sendo usados **2.000 bytes** para dados. Para esta lista, consideramos que estão armazenados na tabela de alunos apenas aqueles que estão atualmente matriculados, e as matrículas se referem apenas aos últimos 5 anos.

Algumas das Estatísticas básicas para essa base de dados são mostradas nas tabelas 1, 2 e 3. Nos exercícios seguintes, sempre complete a tabela 3 com os dados referentes aos índices que cada exercício requer.

Nome	RowSIZE	CARD	MaxCARD	NPags	NExtents
Alunos	152	80.000	80.000	7.273	728
Matriculas	29	6.400.000	6.400.000	110.334	11.034
...		...			

Tabela 1: Métricas para Relações

Relação	Atrib	ColSIZE	ColCARD	Distr	Low2Key	high2Key
Alunos	NUSP	10	80.000	2	'0011223344'	'9988776655'
Alunos	Nome	40	79.500	2	'Abraão Abraão'	'Zuleica Penultima'
Alunos	NomeMae	40	60.000	2	'Alba Bianco'	'Zuleica Penultima'
Alunos	Idade	2	50	2	17	80
Alunos	Cidade	30	700	2	'Adamantina'	'Zzy'
Alunos	Curso	30	400	2	'Agrimensura'	'Zoologia'
Matriculas	Disciplina	8	12.000	2	'AAA-1001'	'ZOO-8228'
Matriculas	NUSP	10	80.000	2	'0011223344'	'9988776655'
Matriculas	Frequencia	2	100	2	1	99
Matriculas	PrimeiraAval	2	100	2	0,1	9,9
Matriculas	SegundaAval	2	100	2	0,1	9,9
Matriculas	Semestre	1	2	2	1	2
Matriculas	Ano	4	77	2	1935	2010
			

Tabela 2: Métricas para Atributos

Relação	Índice	Type	H	NLeafs	Key1Card	KeyFullCard	ClusterRatio
Alunos	AlunosPK	B-tree	3	1.026	80.000	80.000	0
Matriculas	MatriculasPK	B-tree	4	142.223	80.000	6.400.000	0
		

Tabela 3: Métricas para Índices

Exercício 1) Considere que o seguinte índice foi criado:

```
CREATE INDEX UNIQUE AlunoIDXNomeNomeMae ON Alunos(Nome, NomeMae, Idade);
```

Suponha que é emitido o seguinte comando SQL:

```
SELECT NUSP, Nome, Idade
FROM Alunos
WHERE Nome = 'jose da silva' AND
NomeMae='ana da silva' AND
Idade=20;
```

Questão 1.1: Indique quais são os atributos indexáveis para quais índices dessa consulta.

Questão 1.2: Indique quais são os predicados de screening dessa consulta.

Questão 1.3: Indique os índices que podem ser usados nessa consulta.

Questão 1.4: Indique todos os métodos de acesso que podem ser usados nessa consulta, com suas respectivas variantes.

Questão 1.5: Faça a previsão de custo dos métodos de acesso que podem ser usados e indique o que deve ser escolhido. Tente explicar porque esse método é melhor que os demais.

Exercício 2) Considere que o seguinte índice foi criado:

```
CREATE INDEX UNIQUE AlunoIDXNomeNomeMae ON Alunos(Nome, NomeMae, Idade);
```

Suponha que é emitido o seguinte comando SQL:

```
SELECT Nome, Idade
FROM Alunos
WHERE Nome = 'jose da silva' AND
       Idade=20 AND
       NomeMae='ana da silva';
```

Questão 2.1: O que muda em relação ao exercício anterior?

Questão 2.2: Quais são agora os atributos indexáveis para quais índices dessa consulta?

Questão 2.3: Quais são agora os predicados de screening dessa consulta?

Questão 2.4: Quais são agora os índices que podem ser usados nessa consulta?

Questão 2.5: Indique quais são agora os métodos de acesso que podem ser usados nessa consulta, com suas respectivas variantes.

Questão 2.6: Faça a previsão de custo dos métodos de acesso que mudaram. Tente explicar o que mudou.

Exercício 3) Considere que o seguinte índice foi criado:

```
CREATE INDEX UNIQUE AlunoIDXNomeNomeMae ON Alunos(Nome, NomeMae, Idade);
```

Suponha que é emitido o seguinte comando SQL:

```
SELECT Nome, Idade, Curso
FROM Alunos
WHERE Nome like 'luis' OR Nome like 'luiz';
```

Questão 3.1: Indique quais são os atributos indexáveis para quais índices dessa consulta.

Questão 3.2: Indique quais são os predicados de screening dessa consulta. Porque?

Questão 3.3: Indique os índices que podem ser usados nessa consulta.

Questão 3.4: Indique todos os métodos de acesso que podem ser usados nessa consulta, com suas respectivas variantes.

Questão 3.5: Faça a previsão de custo dos métodos de acesso que podem ser usados e indique o que deve ser escolhido. Tente explicar porque esse método é melhor que os demais.

Exercício 4) Considere que os seguintes índices foram criados:

```
CREATE INDEX AlunoIDXNomeIdade ON Alunos(NUSP, Nome, Idade);
CREATE INDEX AlunoIDXIdadeCurso ON Alunos(NUSP, Idade, Curso);
```

Suponha que é emitido o seguinte comando SQL:

```
SELECT Nome, Idade, Curso
FROM Alunos
WHERE NUSP='1231231231';
```

Questão 4.1: Indique quais são os atributos indexáveis para quais índices dessa consulta.

Questão 4.2: Indique quais são os predicados de screening dessa consulta.

Questão 4.3: Indique os índices que podem ser usados nessa consulta.

Questão 4.4: Indique todos os métodos de acesso que podem ser usados nessa consulta, com suas respectivas variantes.

Questão 4.5: Faça a previsão de custo dos métodos de acesso que podem ser usados e indique o que deve ser escolhido. Tente explicar porque esse método é melhor que os demais.

Suponha agora que o índice `AlunoIDXNomeIdade` foi modificado para
`CREATE INDEX AlunoIDXNomeIdade ON Alunos(NUSP, Nome, Curso, Idade);`

Questão 4.6: O que muda?

Questão 4.7: Faça a previsão de custo dos métodos de acesso que podem ser usados e indique o que deve ser escolhido. Tente explicar porque esse método é melhor que os demais.

Exercício 5) Considere que o seguinte índice foi criado:

```
CREATE INDEX AlunoIDXNUSPCurso ON Alunos(NUSP, Curso)
    WHERE Cidade<>'Sao Paulo';
```

Considere também que 54% dos alunos são efetivamente da cidade de São Paulo, e suponha que são emitidos os seguintes comandos SQL:

```
SELECT NUSP, Nome, Curso
    FROM Alunos
    WHERE NUSP =123212321;
```

```
SELECT NUSP, Nome, Curso
    FROM Alunos
    WHERE NUSP =123212321 AND
    WHERE Cidade=São Paulo;
```

```
SELECT NUSP, Nome, Curso
    FROM Alunos
    WHERE NUSP =123212321 AND
    WHERE Cidade<>São Paulo;
```

```
SELECT NUSP, Nome, Curso
    FROM Alunos
    WHERE NUSP =123212321 AND
    WHERE Cidade=Sao Carlos;
```

Questão 5.1: Indique quais são os atributos indexáveis para quais índices de cada consulta.

Questão 5.2: Indique quais são os predicados de *screening* de cada consulta. Porque?

Questão 5.3: Indique os índices que podem ser usados em cada consulta.

Questão 5.4: Indique todos os métodos de acesso que podem ser usados em cada consulta, com suas respectivas variantes.

Questão 5.5: Faça a previsão de custo dos métodos de acesso que podem ser usados em cada consulta e indique o que deve ser escolhido. Tente explicar porque esse é o melhor método em cada consulta.

Exercício 6) Suponha que é emitido o seguinte comando SQL:

```
SELECT A.Nome, M.Disciplina, COALESCE(SegundaAval, PrimeiraAval,0) Nota
    FROM Alunos A NATURAL JOIN Matriculas M
    WHERE Disciplina='SCC-0243' AND
    Ano=2011 AND
    Semestre=1;
```

Suponha os seguintes casos distintos para a existência de índices, e para cada caso responda todas as questões colocadas:

a) Existe apenas as chaves primárias de cada relação.

b) Além das chaves primárias, existe a seguinte chave:

```
CREATE INDEX MatrDiscipAS ON Matriculas (Ano, Semestre, Disciplina);
```

c) Além das chaves primárias, existe a seguinte chave:

```
CREATE INDEX MatrDiscipAS ON Matriculas (Ano, Semestre, Disciplina, NUSP);
```

Questão 6.1: Indique quais são os atributos indexáveis para quais índices dessa consulta em cada caso.

Questão 6.2: Indique quais são os predicados de screening dessa consulta em cada caso.

Questão 6.3: Indique os índices que podem ser usados nessa consulta em cada caso.

Questão 6.4: Indique todos os métodos de acesso que podem ser usados em cada caso nessa consulta, com suas respectivas variantes.

Questão 6.5: Faça a previsão de custo dos métodos de acesso que podem ser usados em cada caso e indique o que deve ser escolhido. Tente explicar porque esse método é melhor que os demais.

Exercício 7) Considere que o seguinte índice foi criado:

```
CREATE INDEX MatrDiscipAS ON Matriculas (Ano, Semestre, Disciplina, NUSP);
```

Suponha que é emitido o seguinte comando SQL:

```
SELECT M.Disciplina, A. Curso,  
       Min(PrimeiraAval), Max(PrimeiraAval), avg(PrimeiraAval)  
FROM Alunos A NATURAL JOIN Matriculas M  
WHERE Ano=2011 AND Semestre=1  
GROUP BY M.Disciplina, A.Curso;
```

Questão 7.1: Indique quais são os atributos indexáveis para quais índices dessa consulta em cada caso.

Questão 7.2: Indique quais são os predicados de screening dessa consulta em cada caso.

Questão 7.3: Indique os índices que podem ser usados nessa consulta em cada caso.

Questão 7.4: Indique todos os métodos de acesso que podem ser usados em cada caso nessa consulta, com suas respectivas variantes.

Questão 7.5: Faça a previsão de custo dos métodos de acesso que podem ser usados em cada caso e indique o que deve ser escolhido. Tente explicar porque esse método é melhor que os demais.