

Estruturas de Indexação de Dados

Cristina D. Aguiar

Índice

- Estrutura de acesso auxiliar usada para melhorar o desempenho na recuperação de registros
 - Pesquisa
 - restringida a um subconjunto dos registros, em contrapartida à análise do conjunto completo
 - realizada em resposta a certas condições
-

Índice

- Estrutura de dados
 - definida em termos das características do índice: registros de dados
 - Operações
 - Pesquisa/Busca
 - Criação
 - Inserção
 - Remoção
 - Atualização
 - Destruição
 - Carregamento
 - Reescrita
-

Pesquisa/Busca

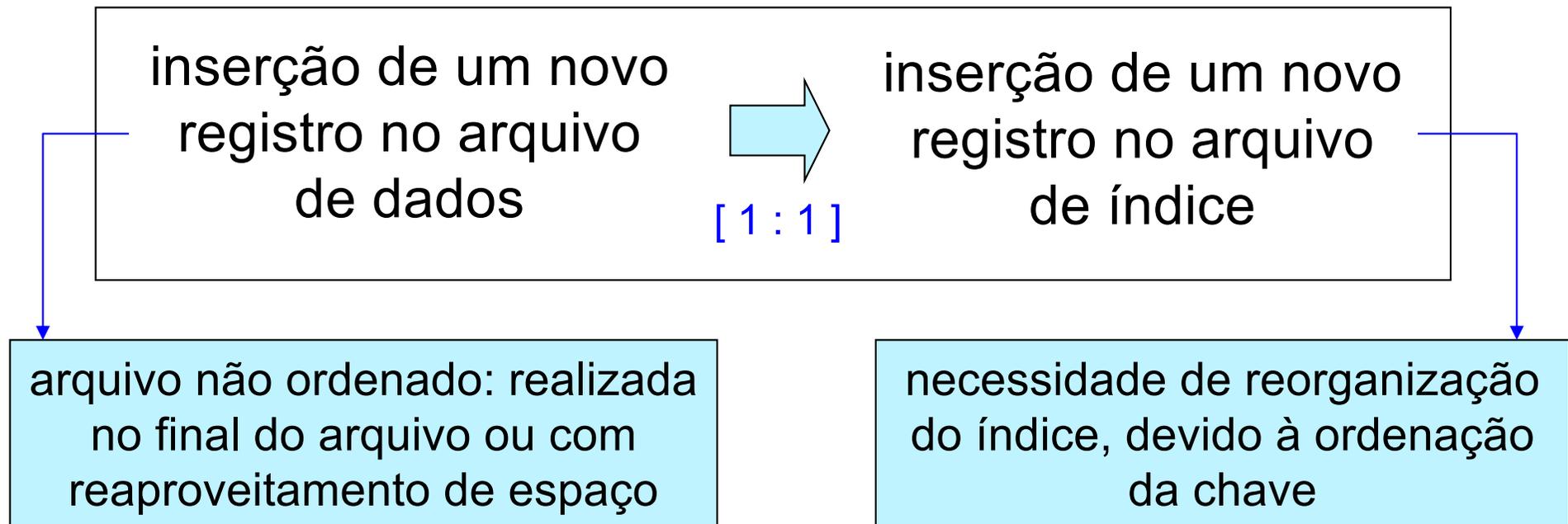
- Baseada na chave de busca
 - encontra a posição da chave no **arquivo de índice**
 - obtém o RRN ou o *byte offset* do registro correspondente à posição encontrada
 - encontra o registro no **arquivo de dados**
 - recupera o registro solicitado do arquivo de dados
-

Criação

- Cria o índice juntamente com a criação do arquivo de dados
 - apenas registro de cabeçalho
 - Cria o índice baseado em um arquivo de dados já existente
 - registro de cabeçalho
 - demais registros (chave de busca + campo de referência), obtidos a partir de uma varredura no arquivo de dados
-

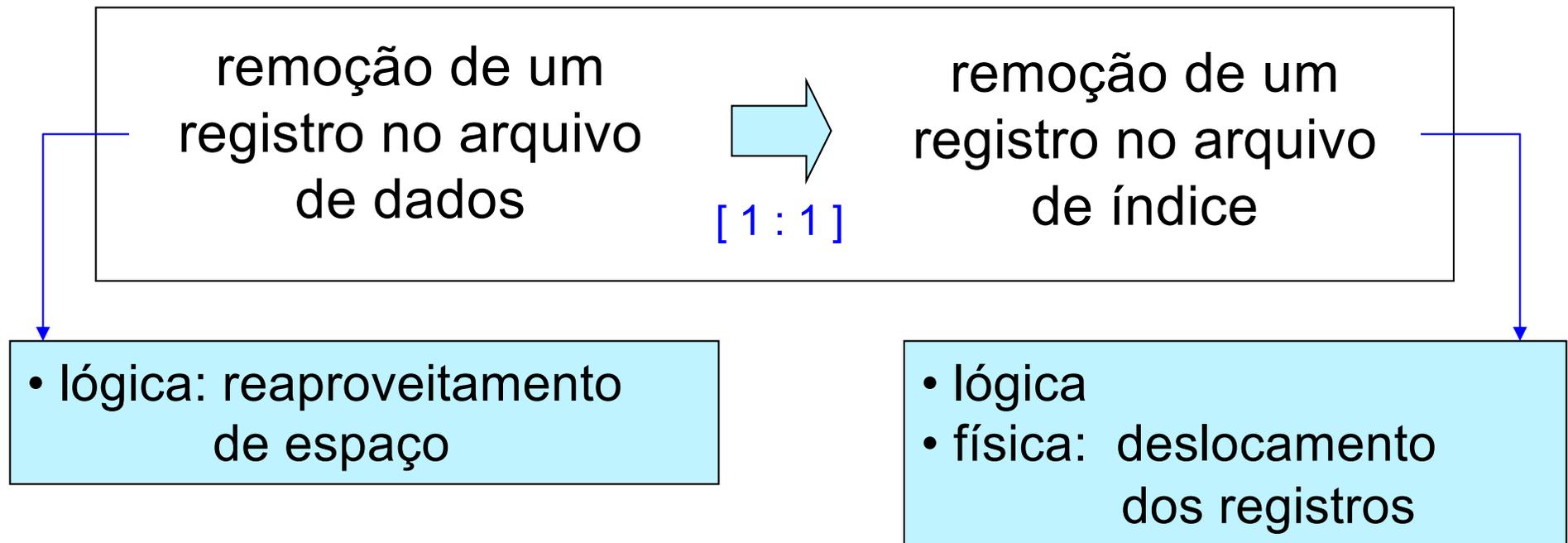
Inserção

- Adiciona registros no índice devido às inserções no arquivo de dados



Remoção

- Remove registros no índice devido às remoções no arquivo de dados



Atualização

- Modifica registros no índice devido às modificações no arquivo de dados
 - Tratamento
 - remoção seguida de inserção ← técnica mais utilizada
ou
 - campo chave: reordenação do índice
 - campo não chave: ajuste do campo de referência se o registro mudar fisicamente no arquivo de dados
-

Destruição

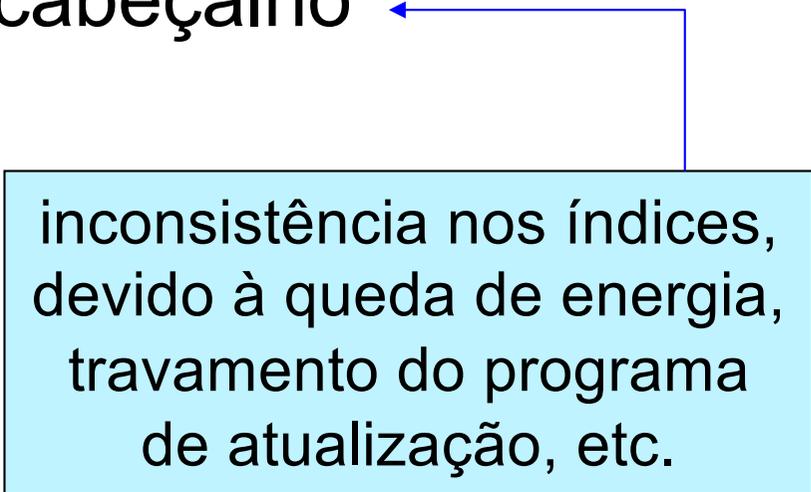
- Exclui o arquivo de índice
- Demais funcionalidades relacionadas
 - realizadas diretamente sobre o arquivo de dados

Carregamento

- Carrega o arquivo de índice na memória principal antes de usá-lo
 - Passos
 - aponta para o primeiro registro do arquivo de índice em disco
 - varre o arquivo de índices sequencialmente
 - cria o índice em memória principal, em geral implementado como um vetor
-

Reescrita

- Atualiza o arquivo de índice em disco com base no arquivo de índice em memória principal, quando necessário
- Informação adicional
 - *status* no registro de cabeçalho
 - verdadeiro/falso



inconsistência nos índices,
devido à queda de energia,
travamento do programa
de atualização, etc.