

Especificação

11/04/2023

1 Informações gerais e objetivos

Há muitos métodos *hashing*, alguns dos quais abordados em aula. Entre os métodos que mantêm o tamanho fixo da tabela *hashing*, o problema de resolver possíveis colisões pode ser resolvido com o uso de ponteiros ou estar baseado em endereçamento aberto. Outras variantes, não vistas em sala, buscam diminuir o tamanho dos ponteiros usados no encadeamento explícito. Estes são argumentos para funções que retornam os endereços do próximo elemento da cadeia de colisões. Para os métodos baseados em endereçamento aberto, estudou-se os efeitos de se usar uma segunda função *hashing* para construir o encadeamento implícito ou de se realocar os registros no momento da inserção de forma melhorar o desempenho durante as consultas.

Esta avaliação tem como objetivo promover o estudo experimental de alguns métodos *hashing* a fim de consolidar o estudo iniciado em aula. Além da necessidade de implementação dos métodos a serem avaliados, os experimentos precisam ser cuidadosamente planejados, evitando viés de comparação. Por exemplo, os dados a serem inseridos na tabela precisam ter suas chaves geradas aleatoriamente. Há ainda a necessidade de repetição para cada experimento, com médias e desvios padrões sendo computados. A final dos experimentos, deve-se ser capaz de identificar se o desempenho esperado dos métodos está condizente com o observado na prática.

2 Objetos de estudo

Os seguintes métodos devem ser incluídos no estudo:

- Encadeamento explícito com o uso de ponteiros usando alocação estática de memória.
- Endereçamento aberto com sondagem linear.
- Endereçamento aberto com duplo *hashing*.
- Outra abordagem *hashing* de livre escolha. Tal escolha deve ser devidamente explicada, o que inclui referências de fontes de estudo usadas para substantiar a escolha.

Espera-se que ao menos as seguintes perguntas sejam respondidas pelo estudo experimental:

1. Quais as diferenças entre os desempenhos dos diferentes métodos e em quais circunstâncias um se sobressai sobre o outro?
2. O desempenho esperado, deduzido teoricamente, reflete o que se observa na prática?
3. Estudos similares anteriores, conduzidos por outros autores, estão em consonância com o que se observou no trabalho? Quais explicações para as diferenças encontradas?

3 Entregáveis

A correção da avaliação será baseada na implementação (peso 50%) e no relatório. O código fonte deverá ser entregue juntamente com o relatório. Plágios, caso detectados, tanto no relatório quanto no código, implicarão nota 0 para todas as equipes envolvidas. Plágios significam cópias entre equipes ou de terceiros.

3.1 Implementação

A linguagem a ser usada para implementação deve ser de uso livre, sem inclusão de bibliotecas proprietárias. Recomenda-se o uso de C, C++ ou Python. Não serão aceitos recursos, presentes em algumas linguagens ou bibliotecas, que contemplem implementação parcial ou total do que se estar querendo avaliar.

A execução do código deve ler um arquivo de entrada em formato texto, contendo uma sequência de inteiros, um por linha. A primeira linha conterá o tamanho m da tabela *hashing* e as demais os $n < m$ valores a serem incluídos na tabela.

3.2 Relatório

A consistência entre o observado na implementação e no relatório estará sendo verificada. O relatório deve ainda conter reflexões sobre as possíveis diferenças encontradas nos experimentos e os resultados teóricos relativos ao desempenho esperado. Apresentação de resultados graficamente, em função do fator de carga ($\alpha = n/m$), é recomendável. Fontes de estudo devem ser devidamente referenciadas no texto do relatório.

O relatório deve conter ao menos as seguintes seções:

- Introdução (máximo uma página).
- Descrição dos métodos objetos de estudo (2-3 páginas).
- Explicação dos experimentos (2-3 páginas).
- Apresentação e discussão dos resultados encontrados (4-5 páginas).
- Conclusão (máximo uma página).

O trabalho pode ser feito em equipes, com no máximo três membros. Espera-se que até dia 18/04/2023 as equipes estejam definidas. Após esta data, equipes não poderão ser modificadas. Nomes não incluídos até esta data em alguma equipe serão interpretados como equipes individuais. A entrega do trabalho (código e relatório) deverá ser feita até dia 14/05/2023.

Tanto as informações sobre as equipes quanto o trabalho deverão ser entregues através do Ava Moodle, na página disciplina, em local indicado.