CE205 Data Structures Week-13

Introduction to File Organization and Processing Sequential File Organization,Direct File Organization Hash Methods

Author: Asst. Prof. Dr. UÄŸur CORUH

Contents

List of Figures

List of Tables

1 CE205 Data Structures

2 Week-13

2.0.1 Introduction to File Organization and Processing Sequential File Organization,Direct File Organization Hash Methods

Download PDF[1],DOCX[2], SLIDE[3], PPTX[4]

2.0.2 Outline-1

- File Organization
  - Sequential File Organization
    * Binary Search
    * Interpolation Search
    * Self-Organizing Sequential Search

---

[1]pandoc_ce205-week-13-direct-sequential-file.tr_doc.pdf
[2]pandoc_ce205-week-13-direct-sequential-file.tr_word.docx
[3]ce205-week-13-direct-sequential-file.tr_slide.pdf
[4]ce205-week-13-direct-sequential-file.tr_slide.pptx

2.0.3 Outline-2

- File Organization
  - Direct File Organization
    * Locating Information
    * Hashing Functions (MD5, HAVAL, SHA1 etc.)
      · Key mod N
      · Key mod P
      · Truncation
      · Folding
      · Squaring
      · Radix Conversion
      · Polynomial Hashing
      · Alphabetic Keys
      · Collisions

2.0.4 Outline-3

- File Organization
  - Direct File Organization
    * Collision Resolution
      · Collision resolution with links
      · Collision resolution without links
      · Static positioning of records
      · Dynamic positioning of records
      · Collision resolution with pseudolinks

2.0.5 Outline-4

- File Organization
  - Direct File Organization
    * Coalesced Hashing
      · EISCH
      · LISCH
      · BEISCH
      · BLISCH
      · REISCH
      · RLISCH
      · EICH
      · LICH

2.0.6 Outline-5

- File Organization
  - Direct File Organization
    * Progressive Overflow
      · Linear Probing
      · Quadratic Probing
    * Double Hashing
    * Use of Buckets
    * Linear Quotient
    * Brentâ€™s Method

2.0.7 Outline-6

- File Organization
  - Direct File Organization
    * Binary Tree
    * Computed Chaining Insertion(CCI)

          ∗ Comparison of Collision Resolution Methods
          ∗ Perfect Hashing
          ∗ SimHash

2.0.8 File Organization

2.0.8.1 Sequential File Organization

- Binary Search
  - https://www.scss.tcd.ie/Owen.Conlan/4d2/4D2-4_File_Sorting_v1.pdf
  - https://www.programiz.com/dsa/binary-search
- Interpolation Search
  - https://www.geeksforgeeks.org/interpolation-search/
- Self-Organizing Sequential Search
  - https://people.csail.mit.edu/rivest/pubs/Riv76a.pdf
  - https://xlinux.nist.gov/dads/HTML/selforganizingSequentialSearch.html
  - https://xlinux.nist.gov/dads/HTML/transposeSeqSearch.html

2.0.9 File Organization

2.0.9.1 Direct File Organization

2.0.9.1.1 Locating Information   Hashing Functions (MD5, HAVAL, SHA1 etc.)

- Key mod N

- Key mod P

- Truncation

- Folding

- Squaring

- Radix Conversion

- Polynomial Hashing

- Alphabetic Keys

- Collisions

- http://www.cs.bilkent.edu.tr/~kdincer/teaching/spring1999/bu-bil212-fo/lectures/pdf-files/bil212-chp6-2.pdf

- https://www.amirajcollege.in/wp-content/uploads/2020/06/3130702-chapter-4-hashing-and-file-structure.pdf

- https://www.cs.bilkent.edu.tr/~kdincer/teaching/spring1999/bu-bil212-fo/lecture_notes.htm

- https://www.cs.otago.ac.nz/cosc242/pdf/L09.pdf

- https://www.cs.otago.ac.nz/cosc242/pdf/L10.pdf

Collision Resolution

- Collision resolution with links
- Collision resolution without links
- Static positioning of records
  - https://www.cs.bilkent.edu.tr/~canf/CS351Fall2010/cs351lecturenotes/week5/index.html
- Dynamic positioning of records
  - https://www.cs.bilkent.edu.tr/~canf/CS351Fall2010/cs351lecturenotes/week5/index.html
- Collision resolution with pseudolinks
  - https://www.cs.bilkent.edu.tr/~canf/CS351Fall2010/cs351lecturenotes/week6/index.html
- http://www.cs.bilkent.edu.tr/~kdincer/teaching/spring1999/bu-bil212-fo/lectures/pdf-files/bil212-chp6-2.pdf

Coalesced Hashing

- EISCH
- LISCH
- BEISCH
- BLISCH
- REISCH
- RLISCH
- EICH
- LICH
- https://www.cs.bilkent.edu.tr/~kdincer/teaching/spring1999/bu-bil212-fo/lectures/pdf-files/bil212-chp6-2.pdf

Progressive Overflow

- Linear Probing
    - https://en.wikipedia.org/wiki/Linear_probing#:~:text=Linear%20probing%20is%20a%20scheme,by%20Gene%2C
- Quadratic Probing
    - https://www.geeksforgeeks.org/quadratic-probing-in-hashing/
- https://www.cs.bilkent.edu.tr/~kdincer/teaching/spring1999/bu-bil212-fo/lectures/pdf-files/bil212-chp6-2.pdf

Double Hashing

- https://www.geeksforgeeks.org/double-hashing/
- https://www.geeksforgeeks.org/hashing-set-3-open-addressing/

Use of Buckets

- https://www.geeksforgeeks.org/file-organization-in-dbms-set-4/

Linear Quotient

- http://www.cs.bilkent.edu.tr/~kdincer/teaching/spring1999/bu-bil212-fo/lectures/pdf-files/bil212-chp6-2.pdf

Brentâ€™s Method

- https://github.com/ncilengir/brent-hashing
- https://cseweb.ucsd.edu//~kube/cls/100/Lectures/lec17.brentsordered/lec17.pdf

Binary Tree

- https://stackoverflow.com/questions/8801898/representing-a-binary-tree-in-a-file
- https://www.geeksforgeeks.org/serialize-deserialize-binary-tree/
- https://www.cs.otago.ac.nz/cosc242/pdf/L12.pdf

Computed Chaining Insertion(CCI)

- https://www.geeksforgeeks.org/c-program-hashing-chaining/

Comparison of Collision Resolution Methods

- https://web.itu.edu.tr/~bkurt/Courses/blg341/lectures_full.pdf

Perfect Hashing

- http://www.cs.otago.ac.nz/cosc242/pdf/L11.pdf

SimHash

- Similar Hash