Lecture 1 Introduction

CS101 Algorithms and Data Structures

Instructor：Dengji Zhao; Yuyao Zhang
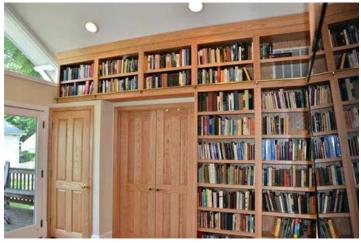
2019-09-10

Outline

- What will we learn from this course?

- How we manage this course?

Definition of Data Structure

- A data structure is a scheme for organizing data in the memory of a computer.

- The way in which the data is organized affects the performance of an algorithm for different tasks.

- 数据结构（data structure）是计算机中存储、组织数据的方式。通常情况下，精心选择的数据结构可以带来最优效率的算法（algorithm）。

Ex1 How to arrange books on the bookshelf?

Ex1 How to arrange books on the bookshelf?

- The following two operations are essential for efficiently arranging your books:

  - **Operation 1:** how to insert new books?

  - **Operation 2:** how to find/access an existing book?

Ex1 How to arrange books on the bookshelf?

- **Method 1:** randomly insert new books.

  - **Operation 1:** how to insert new books?

  Insert the book wherever there is an available space. Nice and easy!

  - **Operation 2:** how to find/access an existing book?

  It depends …

Ex1 How to arrange books on the bookshelf?

- **Method 2:** insert new books according to the alphabets order of the first letter.

  - **Operation 1:** how to insert new books?

  EX: we bought a new book "Algorithm".

  - **Operation 2:** how to find/access an existing book?

  EX: Binary search!

Ex1 How to arrange books on the bookshelf?

- **Discussion 1:** is **Method 2** absolutely better/more efficient than **Method 1**?

Method 1: randomly insert new books.

Method 2: insert new books according to the alphabets order of the first letter.

- **Discussion 2:** how can we further improve **Method 2**?

Ex1 How to arrange books on the bookshelf?

- **Method 3: cluster books according to different topics** (computer science, economics, agriculture, politics…), then insert new books according to the alphabets order of the first letter.

  - **Operation 1:** how to insert new books?

    EX: we bought a new book "Algorithm".

  - **Operation 2:** how to find/access an existing book?

    EX: Binary search for topic first, then binary search for book title.

- ➢ **Discussion 3:** how much space we should preserve for each topic? How many topics is an optimism option?

The efficiency of a method/algorithm highly depends on the organization&amount of the data.

Ex2 How to implement a function PrintN?

- Implement a function named PrintN, when input a positive integer N, print all the positive integer from 1 to N.

| ```c
void PrintN ( int N )
{ int i;
  for ( i=1; i<=N; i++ ) {
      printf( "%d\n" , i);
}
return;
}
``` | ```c
void PrintN ( int N )
{ if (N);
   PrintN( N-1);
   printf( "%d\n" , N);
}
return;
}
``` |
| Loop implementation | Recursive implementation |

- Let N = 100, 1000, 10000, 100000, … …

Ex2 How to implement a function PrintN?

- Implement a function named PrintN, when input a positive integer N, print all the positive integer from 1 to N.

```c
# include <stdio.h>
void PrintN ( int N );
int main ()
{ int N;
  scanf ("%d", &N);
  PrintN(N);
  return 0;
}
```


Loop implementation

Recursive implementation

why?

The efficiency of a method/algorithm depends on the occupation of RAM.

Definition of Data Structure

- **Data structure**, way in which data are stored for efficient search and retrieval.

- Different data structures are suited for different **operations**.

- **Algorithm** is a procedure for solving a mathematical problem in a finite number of steps that frequently involves repetition of an operation.

Abstract Data Type (ADT 抽象数据类型)

- Abstract: The method that we describe the data type, does not depend on the implementations.

  - Not related to the computer that stores the data.

  - Not related to the physical structure that stores the data.

  - Not related to the algorithm and language that implements the operation.

- We only care about "*how to design*" the objective data sets and related operations, not how to "*implement*" a data structure.

EX4 Abstract data type of a $matrix$

array?
structural array? Orthogonal list?

- Data type: $Matrix$
- Objects: a $M \times N$ matrix $\underline{A_{M \times N} = (a_{ij})}$ $(i = 1, \cdots, M; j = 1, \cdots, N)$ is composed by a number of $M \times N$ array of $< a, i, j >$, where a present the value of the matrix element, i present the no. of row, and j present the no. of column.
- Operations: for an arbitrary matrix $A, B, C \in Matrix$, and integers i, j, M, N
- $Matrix\ create\ (int\ M, int\ N)$: return an empty matrix of $M \times N$;
- $int\ GetMaxRow(\ Matrix\ A\ )$: return the number of rows;
- $int\ GetMaxCol(\ Matrix\ A\ )$: return the number of columns;
- $\underline{ElementType}\ GetEntry(\ Matrix\ A, int\ i, int\ j)$: return the element of matrix A in row i, column j;
- $Matrix\ Add\ (Matrix\ A,\ Matrix\ B)$: if the dimension of matrix A and B are the same, return matrix $\underline{C = A + B}$, otherwise error;
- $Matrix\ multiply\ (Matrix\ A,\ Matrix\ B)$: if the number of columns of matrix A is equals to the number of rows of matrix B, return matrix $C = AB$, otherwise return error;
- ……

The elements are added in order of rows or columns? C,C++,Python,…?