Data Structures
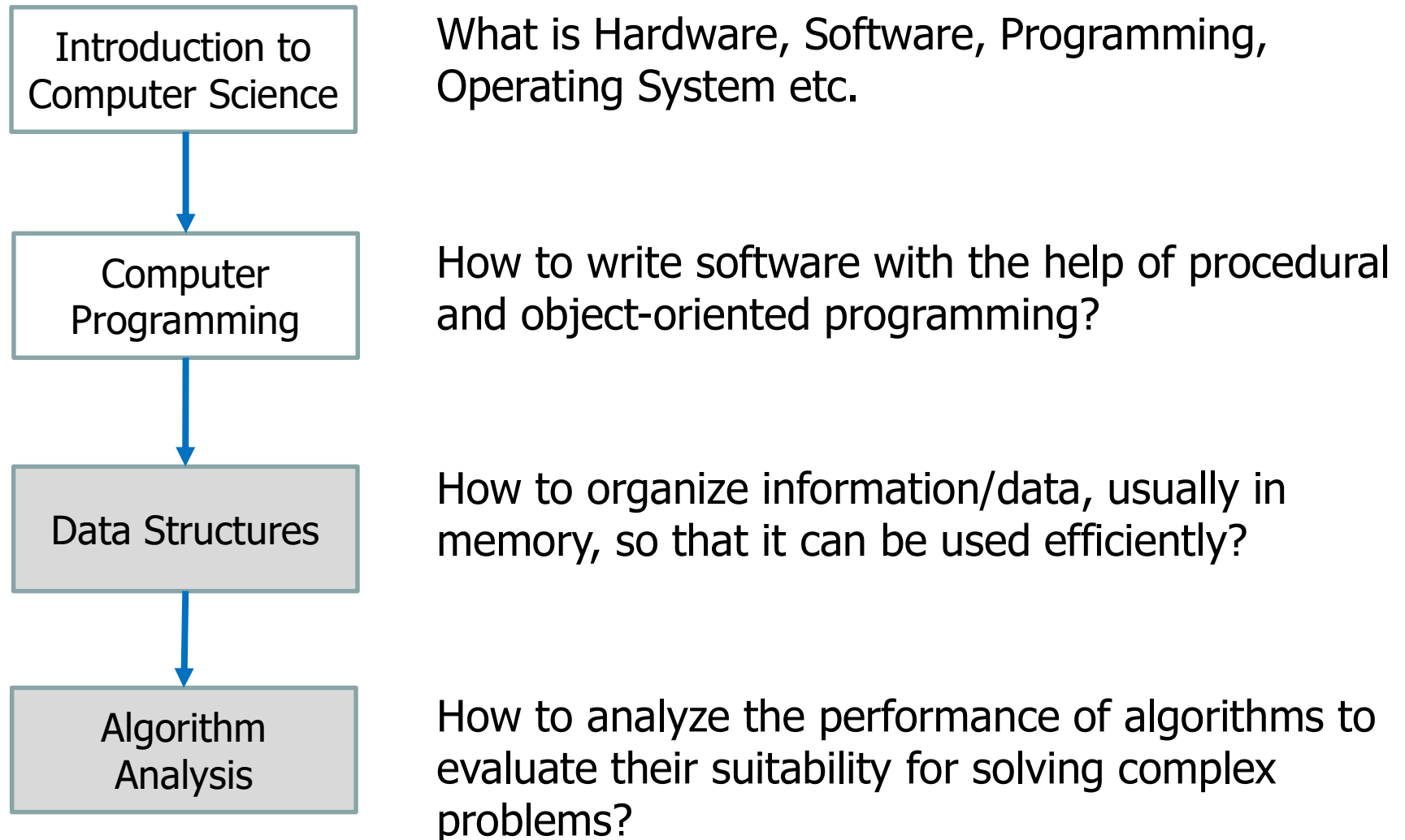## Fall 2021

Dr. Adnan Tariq

muhammad.adnan@nu.edu.pk

National University of Computer and Emerging Sciences

General Overview

Introduction to Computer Science	What is Hardware, Software, Programming, Operating System etc.
↓	
Computer Programming	How to write software with the help of procedural and object-oriented programming?
↓	
Data Structures	How to organize information/data, usually in memory, so that it can be used efficiently?
↓	
Algorithm Analysis	How to analyze the performance of algorithms to evaluate their suitability for solving complex problems?

What is a Data Structure?

- In a general sense, any representation that is used for storing information is a data structure
  - Example: integers, structures, classes, arrays, etc.

- More typically, a data structure provides a way for organizing a collection of data items
  - Storing, accessing and modifying data items

- Most data structures have associated algorithms to perform operations on data
  - Search, insert, remove etc.

Costs and Benefits

- Each data structure requires
  - Space for storing data items
  - Time to perform each basic operation
  - Programming efforts

- Cost of a data structure
  - Time and space resources in consumes

- Choice of data structure depends on many factors
  - Type of data
  - Frequency with which various data operations are applied

- Hard to define data structure that performs better in all situations
  - Time and space tradeoff

Goals of This Course

- Learn the commonly used data structures
  - Form a programmer's basic data structure "toolkit"

- Case studies of data structures

- Examine the costs and benefits of every data structure

Course Content

- Simple and abstract data types
- Background: Algorithm Analysis
- Arrays
- Searching and sorting techniques
- Linked Lists
- Queues
- Stacks
- Trees: Definitions and terminology

Mid term

- -

- Trees: Binary Search Trees, AVL Trees
- Graphs
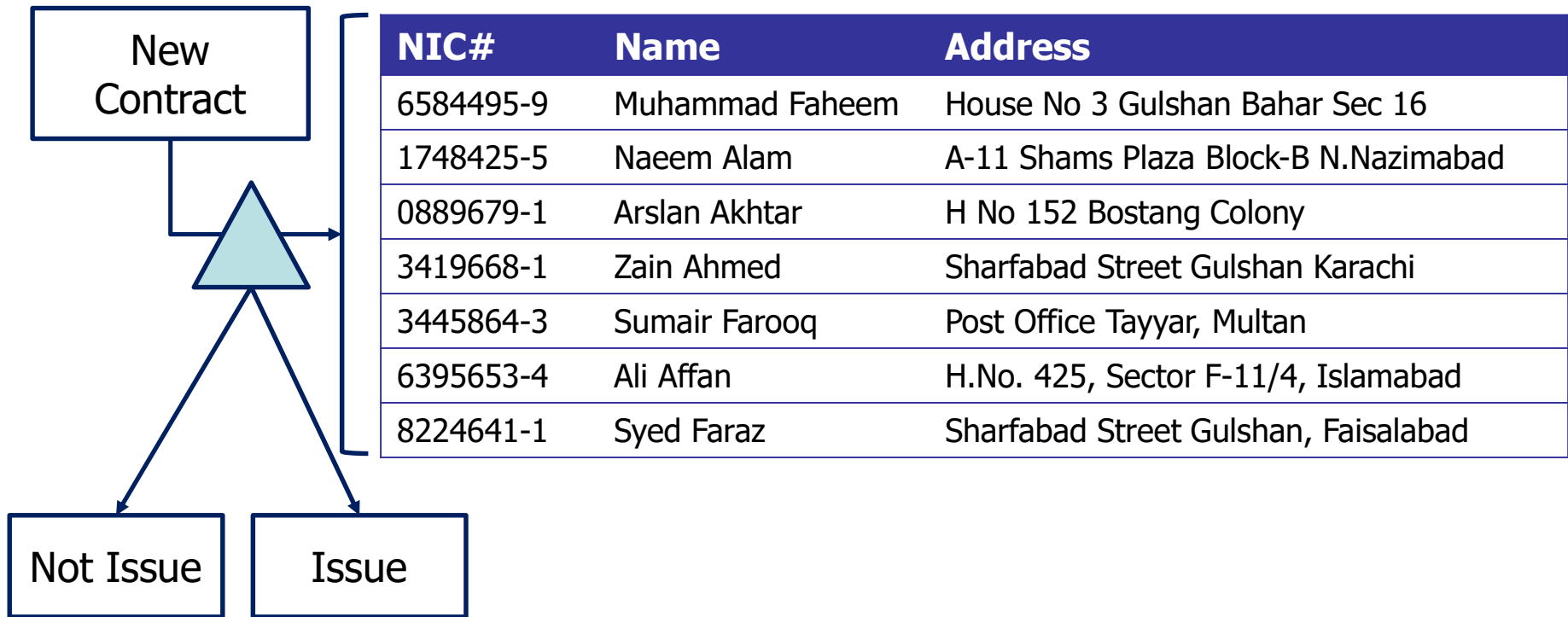- Hashing
- Heap
- B+ Trees

Motivational Example

- A cellular service company provides contract to its 10 million users
- Due to new security enforcements, the company wants to prevent issuing of multiple contracts to users

- Method of Detecting Multiple Contracts
  - Before issuing a new contract to user
  - First search the id of user in existing contracts database
  - In case of failure → issue a new contract
  - In case of success → do not issue a new contract to user

Example: Linear Array Data Structure (1)

NIC#	Name	Address
6584495-9	Muhammad Faheem	House No 3 Gulshan Bahar Sec 16
1748425-5	Naeem Alam	A-11 Shams Plaza Block-B N.Nazimabad
0889679-1	Arslan Akhtar	H No 152 Bostang Colony
3419668-1	Zain Ahmed	Sharfabad Street Gulshan Karachi
3445864-3	Sumair Farooq	Post Office Tayyar, Multan
6395653-4	Ali Affan	H.No. 425, Sector F-11/4, Islamabad
8224641-1	Syed Faraz	Sharfabad Street Gulshan, Faisalabad

- Linear Array (with 10 million entries)
  - 3 arrays (NIC, Name, Address)
  - Structured array
  - Class's object array

Example: Linear Array Data Structure (2)

```
New
Contract
```

NIC#	Name	Address
6584495-9	Muhammad Faheem	House No 3 Gulshan Bahar Sec 16
1748425-5	Naeem Alam	A-11 Shams Plaza Block-B N.Nazimabad
0889679-1	Arslan Akhtar	H No 152 Bostang Colony
3419668-1	Zain Ahmed	Sharfabad Street Gulshan Karachi
3445864-3	Sumair Farooq	Post Office Tayyar, Multan
6395653-4	Ali Affan	H.No. 425, Sector F-11/4, Islamabad
8224641-1	Syed Faraz	Sharfabad Street Gulshan, Faisalabad

```
Not Issue          Issue
```
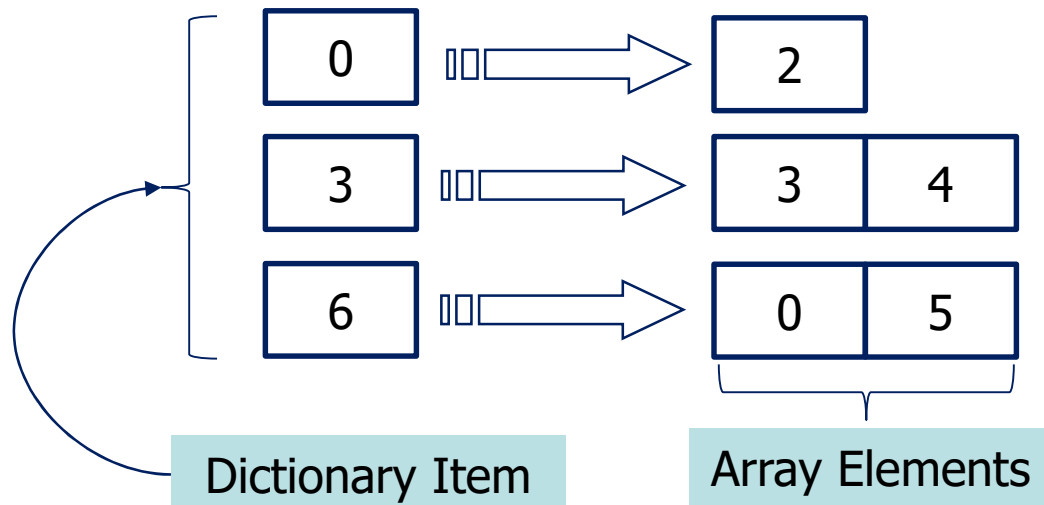
- Any disadvantage of linear array data structure?
- How to improve?

Example: Improved Data Structure

- Create a dictionary data structure
  - Group similar records together
  - Similarity in terms of first digit of NIC number
  - Add a dictionary entry for each distinct digit (0 – 9)

- Example: 3419668-1, 3445864-3, 1748425-5.
  - 3 and 1 are dictionary entries

- Existing contracts are searched in two steps
  1. Search the dictionary entry (i.e., group searching)
  2. Search contract within the group (i.e., with the same NIC digit)
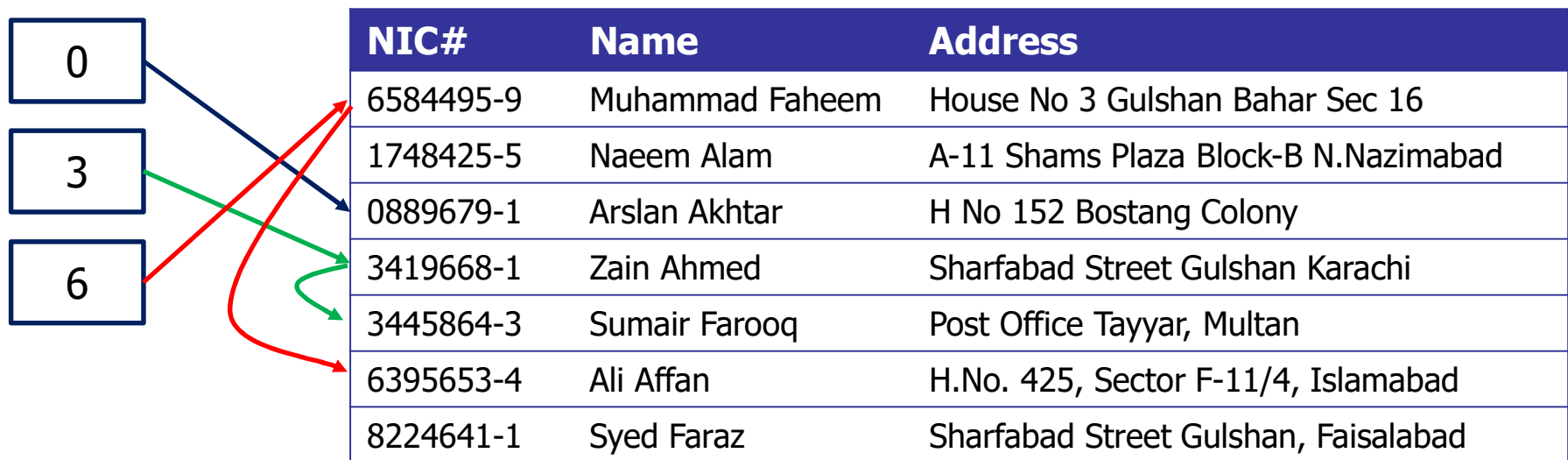
- How much improvement w.r.t. linear array?

Introduction

Example: Dictionary Data Structure (1)

NIC#	Name	Address
6584495-9	Muhammad Faheem	House No 3 Gulshan Bahar Sec 16
1748425-5	Naeem Alam	A-11 Shams Plaza Block-B N.Nazimabad
0889679-1	Arslan Akhtar	H No 152 Bostang Colony
3419668-1	Zain Ahmed	Sharfabad Street Gulshan Karachi
3445864-3	Sumair Farooq	Post Office Tayyar, Multan
6395653-4	Ali Affan	H.No. 425, Sector F-11/4, Islamabad
8224641-1	Syed Faraz	Sharfabad Street Gulshan, Faisalabad

Dictionary Item   Array Elements

Example: Dictionary Data Structure (2)

- Another possibility
  - Maintain pointers with structures (or records)
  - Non NULL pointer indicates presence of next record

0

3

6

NIC#	Name	Address
6584495-9	Muhammad Faheem	House No 3 Gulshan Bahar Sec 16
1748425-5	Naeem Alam	A-11 Shams Plaza Block-B N.Nazimabad
0889679-1	Arslan Akhtar	H No 152 Bostang Colony
3419668-1	Zain Ahmed	Sharfabad Street Gulshan Karachi
3445864-3	Sumair Farooq	Post Office Tayyar, Multan
6395653-4	Ali Affan	H.No. 425, Sector F-11/4, Islamabad
8224641-1	Syed Faraz	Sharfabad Street Gulshan, Faisalabad

Some Rules

- Never ever miss a class
  - No retakes (except for Final/Mid Exam*)
- Never use mobile phone in the class
- Above all, whatever you do, please do not disturb others

* Conditional: as per university policy

Dishonesty and Plagiarism

- Any kind of cheating will be considered serious offense

- Plagiarism in project or midterm/ final exam may result in F grade in the course.

- Plagiarism in an assignment will result in zero marks in the whole assignments category.
  - All parties involve in cheating will get zero marks

Attendance Policy

- Student arriving after the attendance will be marked "absent"
  - No excuse to students arriving immediately after the attendance

- 80% attendance is mandatory for participating in the final exam

- Students are not allowed to switch sections

Tentative Evaluation Breakdown

Evaluation Name	Weightage
Lab	15
Assignment (3)	15
Project	10
Quiz (>4)	05
Midterm	15
Final	40

- Absolute grading …. ☺
- Lab attendance is mandatory!
  - Irrespective to course code 218, 201!
  - Hard to get good marks in assignments without lab

Learning Management System

- Google Classroom
  - Announcements
  - Material/Slides
  - Assignments

- How to access?
  - Code: **6ws5cfu**
  - https://classroom.google.com/c/MzIwMTM3MzA0NTgw?cjc=6ws5cfu

Course & Reference Books

Books in Library

- D.S. Malik, Data Structures Using C++, 8th Edition
- Mark Allen Weiss, Data Structures and Algorithm Analysis in C++, 4th Edition
- James Robergé, Stefan Brandle, David Whittington, A laboratory course in C++ Data structures, 2nd Edition

PDF Books

- Mark Allen Weiss, Data Structures and Problem-Solving Using C++, 2nd Edition
- D.S.Malik, C++ programming: Program design including data structures, 5th Edition
- Clifford A. Shaffer, Data Structures and Algorithm Analysis, Edition 3.

Any Question So Far?