### **Java Collections Framework**

### Ameya Joshi

ameya.joshi\_official@outlook.com +91 9850676160

1

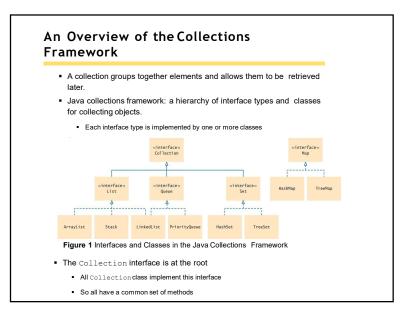
### **What is Collections Framework**

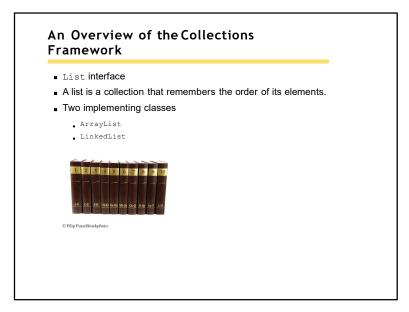
The Java collections framework standardizes the way in which groups of objects are handled by your programs.

It is a well-designed set of interfaces and classes for storing and manipulating groups of data as a single unit, a  $\it collection$ .

The framework provides a convenient API to many of the abstract data types familiar from computer science data structure curriculum: maps, sets, lists, trees, arrays, hashtables, and other collections.

Because of their object-oriented design, the Java classes in the Collections Framework encapsulate both the data structures and the algorithms associated with these abstractions.





# An Overview of the Collections Framework

- Set interface
- A set is an **unordered** collection of **unique** elements.
- Arranges its elements so that finding, adding, and removing elements is more efficient.
- Two mechanisms to do this
  - hash tables
  - binary search trees



5

# An Overview of the Collections Framework

- Stack
  - Remembers the order of elements
  - But you can only add and remove at the top



# An Overview of the Collections Framework

- Queue
  - Add items to one end (the tail) and remove them from the other end (the head)
- A queue of people

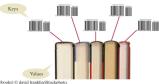


- A priority queue
  - an unordered collection
  - has an efficient operation for removing the element with the highest priority

7

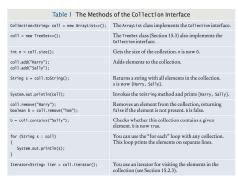
# An Overview of the Collections Framework

- Map
  - Keeps associations between key and value objects.
  - Every key in the map has an associated value.
  - The map stores the keys, values, and the associations between them.



## An Overview of the Collections Framework

 $\blacksquare$  Every class that implements the  ${\tt Collection}$  interface has these methods.



9

### Self Check

A gradebook application stores a collection of quizzes. Should it use a list or a set?

**Answer:** A list is a better choice because the application will want to retain the order in which the quizzes were given.

A student information system stores a collection of student records for a university. Should it use a list or a set?

**Answer:** A set is a better choice. There is no intrinsically useful ordering for the students. For example, the registrar's office has little use for a list of all students by their roll No.. By storing them in a set, adding, removing, and finding students can be efficient.

11

### Self Check

Why is a queue of books a better choice than a stack for organizing your required reading?

**Answer:** With a stack, you would always read the latest required reading, and you might never get to the oldest readings.

As you can see from Figure 1, the Java collections framework does not consider a map a collection. Give a reason for this decision.

**Answer:** A collection stores elements, but a map stores associations between elements.

13

### The Diamond Syntax

 Convenient syntax enhancement for array lists and other generic classes.

.

■ You can write:

ArrayList<String> names = new ArrayList<>();

instead of:

ArrayList<String> names = new ArrayList<String>();

- This shortcut is called the "diamond syntax" because the empty brackets <> look like a diamond shape.
- This chapter, and the following chapters, will use the diamond syntax for generic classes.

### **Linked Lists**

- A data structure used for collecting a sequence of objects:
  - Allows efficient addition and removal of elements in the middle of the sequence.
- A linked list consists of a number of nodes;
  - Each node has a reference to the next node.
- A node is an object that stores an element and references to the neighboring nodes.
- Each node in a linked list is connected to the neighboring nodes.



© andrea laurita/iStockphoto.

15

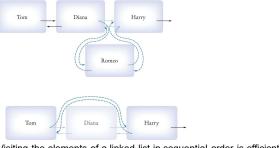
### Linked Lists

- Adding and removing elements in the middle of a linked list is efficient.
- Visiting the elements of a linked list in sequential order is efficient.
- Random access is **not** efficient.



### **Linked Lists**

- When inserting or removing a node:
  - Only the neighboring node references need to be updated



Visiting the elements of a linked list in sequential order is efficient.

- Random access is not efficient.

17

### **Linked Lists**

- When to use a linked list:
  - You are concerned about the efficiency of inserting or removing elements
  - You rarely need element access in random order

# The LinkedList Class of the Java Collections Framework

- Generic class
  - Specify type of elements in angle brackets: LinkedList<Product>
- Package: java.util
- LinkedList has the methods of the Collection interface.
- Some additional LinkedList methods:

Table 2 Working with Linked Lists		
LinkedList <string> list = new LinkedList&lt;&gt;();</string>	An empty list.	
list.addLast("Harry");	Adds an element to the end of the list. Same as add.	
<pre>list.addFirst("Sally");</pre>	Adds an element to the beginning of the list. 1ist is now [Sally, Harry].	
list.getFirst();	Gets the element stored at the beginning of the list; here "Sally".	
list.getLast();	Gets the element stored at the end of the list; here "Harry".	
<pre>String removed = list.removeFirst();</pre>	Removes the first element of the list and returns it. removed is "Sally" and list is [Harry]. Use removeLast to remove the last element.	
ListIterator <string> iter = list.listIterator()</string>	Provides an iterator for visiting all list elements (see Table 3 on page $684). \\$	

19

### List Iterator

- $_{\hbox{\scriptsize \blacksquare}}$  Use a list iterator to access elements inside a linked list.
- Encapsulates a position anywhere inside the linked list.
- Think of an iterator as pointing between two elements:

Analogy: like the cursor in a word processor points between two characters

■ To get a list iterator, use the listIterator method of the LinkedList class.

LinkedList<String> employeeNames = . . .; ListIterator<String> iterator = employeeNames.listIterator();

Also a generic type.

### List Iterator

- Initially points before the first element.
- $_{\bullet}$  Move the position with next method:

```
if (iterator.hasNext())
{
   iterator.next();
}
```

- $\bullet$  The  $\mathtt{next}$  method returns the element that the iterator is passing.
- $\blacksquare$  The return type of the  $\mathtt{next}$  method matches the list iterator's type parameter.

21

### List Iterator

■ To traverse all elements in a linked list of strings:

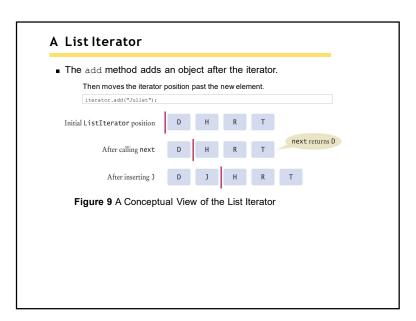
```
while (iterator.hasNext())
{
   String name = iterator.next();
   Do something with name
```

■ To use the "for each" loop:

```
for (String name : employeeNames) {
    Do something with name }
```

# The nodes of the LinkedList class store two links: One to the next element One to the previous one Called a doubly-linked list To move the list position backwards, use: hasPrevious previous

23



### List Iterator

■ The remove method:

Removes object that was returned by the last call to next or previous

• To remove all names that fulfill a certain condition:

```
while (iterator.hasNext())
{
    string name = iterator.next();
    if (condition is fulfilled for name)
    iterator.remove();
}
```

■ Be careful when calling remove:

It can be called only **once** after calling next or previous You cannot call it immediately after a call to add

If you call it improperly, it throws an  ${\tt IllegalStateException}$ 

25

### List Iterator

- $_{\blacksquare}$  ListIterator interface extends Iterator interface.
- Methods of the Iterator and ListIterator Interfaces

Table 3 Methods of the Iterator and ListIterator Interfaces		
String s = iter.next(); Assume that iter points to the beginning of the list [\$a11y] before calling next. After the call, s is "\$a11y" and the iterator points to the end.		
<pre>iter.previous();</pre>		
iter.hasNext() Returns false because the iterator is at the end of the collection		
if (iter.hasPrevious()) {	ing	
iter.add("Diana"); Adds an element before the iterator position (ListIterator only The list is now [Diana, Juliet].	·).	
tem:tem:tem:tem:tem:tem:tem:tem:tem:tem:	ne	

Do linked lists take more storage space than arrays of the same size?

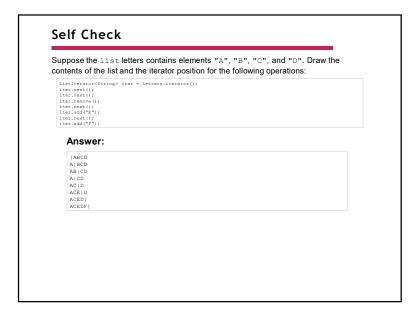
**Answer:** Yes, for two reasons. A linked list needs to store the neighboring node references, which are not needed in an arry. Moreover, there is some overhead for storing an object. In a linked list, each node is a separate object that incurs this overhead, whereas an array is a single object.

27

### Self Check

Why don't we need iterators with arrays?

**Answer:** We can simply access each array element with an integer index.



29

# Self Check Write a loop that removes all strings with length less than four from a linked list of strings called words. Answer: ListIterator<String> iter = words.iterator(); while (iter.hasRext()); { String str = iter.next(); if (str.length() < 4) { iter.remove(); } }

Write a loop that prints every second element of a linked list of strings called words.

### Answer

```
ListIterator<String> iter = words.iterator();
while (iter.hasNext())
{
    System.out.println(iter.next());
    if (iter.hasNext())
    {
        iter.next(); // Skip the next element
    }
}
```

31

### Sets

- A set organizes its values in an order that is optimized for efficiency.
- May not be the order in which you add elements.
- Inserting and removing elements is more efficient with a set than with a list.

### Sets

- The Set interface has the same methods as the Collection interface.
- A set does not admit duplicates.
- Two implementing classes

HashSet based on hash table
TreeSet based on binary search tree

A Set implementation arranges the elements so that it can locate them quickly.

33

### Sets

■ In a hash table

Set elements are grouped into smaller collections of elements that share the same characteristic.

Grouped by an integer hash code that is computed from the element.

- Elements in a hash table must implement the method hashCode.
- $_{\blacksquare}$  Must have a properly defined  ${\tt equals}$  method.
- ${\color{red} \bullet}$  You can form hash sets holding objects of type  ${\tt String},$

Integer, Double, Point, Rectangle,  $\ensuremath{\text{or}}$  Color.

HashSet<String>, HashSet<Rectangle>, or a
HashSet<HashSet<Integer>>

On this shelf, books of the same color are grouped together.
 Similarly, in a hash table, objects with the same hash code are placed in the same group.

Alfredo Ragazzoni/iStockphot

### Sets

■ In a TreeSet

Elements are kept in sorted order



© Volkan Ersoy/iStockphoto.

- Elements are stored in nodes.
- The nodes are arranged in a tree shape,

Not in a linear sequence

You can form tree sets for any class that implements the Comparable interface:

 $\textbf{Example:} \, \texttt{String} \, \textbf{or} \, \texttt{Integer.}$ 

35

### Sets

■ Use a TreeSet if you want to visit the set's elements in sorted order

Otherwise choose a HashSet.

It is a bit more efficient — if the hash function is well chosen.

### Sets

 $\blacksquare$  Store the reference to a TreeSet or HashSet in a

Set<String> variable:

Set<String> names = new HashSet<>();
or
Set<String> names = new TreeSet<>();

■ After constructing the collection object:

The implementation no longer matters Only the interface is important

37

### **Working with Sets**

Adding and removing elements:

names.add("Romeo");
names.remove("Juliet");

Sets don't have duplicates.

Adding a duplicate is ignored.

- Attempting to remove an element that isn't in the set is ignored.
- The contains method tests whether an element is contained in the set:

if (names.contains("Juliet")) . . .

The  ${\tt contains}$  method uses the  ${\tt equals}$  method of the element type

### **Working with Sets**

- To process all elements in the set, get an iterator.
- A set iterator visits the elements in the order in which the set implementation keeps them.

```
Iterator<String> iter = names.iterator();
while (iter.hasNext())
{
   String name = iter.next();
   Do something with name
}
}
```

■ You can also use the "for each" loop

```
for (String name : names) {
    Do something with name }
}
```

- You cannot add an element to a set at an iterator position A set is unordered.
- You can remove an element at an iterator position.
- The iterator interface as no previous method.

39

Arrays and lists remember the order in which you added elements; sets do not. Why would you want to use a set instead of an array or list?

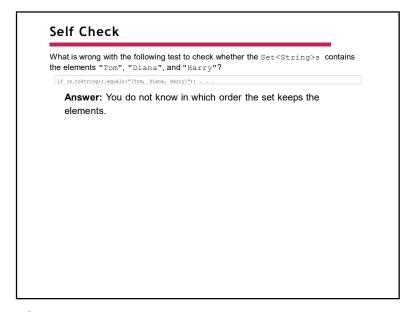
**Answer:** Adding and removing elements as well as testing for membership is more efficient with sets.

41

### Self Check

Why are set iterators different from list iterators?

**Answer:** Sets do not have an ordering, so it doesn't make sense to add an element at a particular iterator position, or to traverse a set backward.



43

```
Self Check

How can you correctly implement the test of Self Check 12?

Answer: Here is one possibility:

if (s.size() == 3 && s.contains("Tom") && s.contains("Diana") && s.contains("Harry")) && s.contains("Harry"))
```

# Self Check Write a loop that prints all elements that are in both Set<String>s and Set<String>t. Answer: for (String str : s) { if (t.contains(str)) { System.out.println(str); } }

45

### **Self Check**

Suppose you changed line of the  ${\tt SpellCheck}$  program to use a {\tt TreeSet} instead of a {\tt HashSet}. How would the output change?

Answer: The words would be listed in sorted order.

### Maps

- A map allows you to associate elements from a key set with elements from a value collection.
- Use a map when you want to look up objects by using a key.

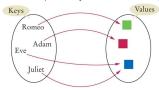


Figure 10 A Map

■ Two implementations of the Map interface:

HashMap TreeMap

• Store the reference to the map object in a Map reference:

Map<String, Color> favoriteColors = new HashMap<>(

47

### Maps

- Use the put method to add an association:
  - favoriteColors.put("Juliet", Color.RED);
- You can change the value of an existing association by calling put again:

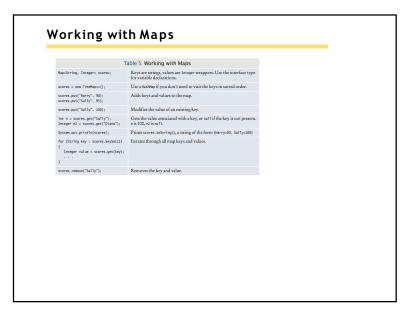
favoriteColors.put("Juliet", Color.BLUE);

 $\blacksquare$  The get method returns the value associated with a key:

Color julietsFavoriteColor = favoriteColors.get("Juliet");

- If you ask for a key that isn't associated with any values, the get method returns null.
- $\blacksquare$  To remove an association, call the  ${\tt remove}$  method with the key:

favoriteColors.remove("Juliet");



49

### Мар

- Sometimes you want to enumerate all keys in a map.
- $_{\bullet}$  The  ${\tt keySet}$  method yields the set of keys.
- Ask the key set for an iterator and get all keys.
- For each key, you can find the associated value with the get method.
- To print all key/value pairs in a map m:

```
Set<String> keySet = m.keySet();
for (String key : keySet)
{
   Color value = m.get(key);
   System.out.println(key + "->" + value);
}
}
```

What is the difference between a set and a map?

**Answer:** A set stores elements. A map stores associations between keys and values.

51

### Self Check

Why is the collection of the keys of a map a set and not a list?

**Answer:** The ordering does not matter, and you cannot have duplicates

Why is the collection of the values of a map not a set?

Answer: Because it might have duplicates.

53

### Self Check

Suppose you want to track how many times each word occurs in a document. Declare a suitable map variable.

Answer: Map<String, Integer> wordFrequency;

What is a Map<String, HashSet<String>>? Give a possible use for such a structure

Answer: It associates strings with sets of strings. One application would be a thesaurus that lists synonyms for a given word. For example, the key "improve" might have as its value the set ["ameliorate", "better", "enhance", "enrich", "perfect", "refine"].

55

### Java 8 Note

■ Updating Map Entries used to be tedious

Integer count = frequencies.get(word); // Get the old frequency count // if there was none, put l; otherwise, increment the count if (count == null) (count == 1; ) else (count == count = 1; ) frequencies.put(word, count);

- New merge method in Map interface
- Specify:

Key

Value to be used if key not yet present

Function to compute the updated value if key is present

Replace code above with single line using lambda expression:

frequencies.merge(word, 1, (oldValue, value) -> oldValue + value);

### **Choosing a Collection**

- 1. Determine how you access the values.
- 2. Determine the element types or key/value types.
- 3. Determine whether element or key order matters.
- 4. For a collection, determine which operations must be efficient.
- 5. For hash sets and maps, decide whether you need to implement the hashCode and equals methods.
- 6. If you use a tree, decide whether to supply a comparator.

57

### **Hash Functions**

- You may need to implement a hash function for your own classes.
- A hash function: a function that computes an integer value, the hash code, from an object in such a way that different objects are likely to yield different hash codes.
- Object class has a hashCode method
   you need to override it to use your class in a hash table
- A collision: two or more objects have the same hash code.
- $\blacksquare$  The method used by the  ${\tt String}$  class to compute the hash code.

```
final int HASH_MULTIPLIER = 31;
int h = 0;
for (int i = 0; i < s.length(); i++)
{
    h = HASH_MULTIPLIER * h + s.charAt(i);
}</pre>
```

 $\blacksquare$  This produces different hash codes for "tea" and "eat".

### **Hash Functions**

A good hash function produces different hash values for each object so that they are scattered about in a hash table.

59

### **Hash Functions**

- Override hashCode methods in your own classes by combining the hash codes for the instance variables.
- A hash function for a Country class:

```
public class Country
{
   public int hashCode()
   {
     int h1 = name.hashCode();
     int h2 = new Double(area).hashCode();
     final int HASH_MULTIPLIER = 29;
     int h = HASH_MULTIPLIER * h1 + h2;
     return h;
   }
}
```

- Easier to use Objects.hash() method
  - Takes hashcode of all arguments and multiplies them:

```
public int hashCode()
{
   return Objects.hash(name, area);
}
```

 $\bullet$  A class's <code>hashCode</code> method must be compatible with its <code>equals</code> method.

### Stacks

• A stack lets you insert and remove elements only at one end:

Called the top of the stack.

Removes items in the opposite order than they were added Last-in, first-out or LIFO order

- Add and remove methods are called push and pop.
- Example

```
Stack<String> s = new Stack<>();
s.push("A"); s.push("B"); s.push("C");
while (s.size() > 0)
{
    System.out.print(s.pop() + " "); // Prints C B A
}
```

■ The last pancake that has been added to this stack will be the first one that is consumed.



61

### Stacks

- Many applications for stacks in computer science.
- Consider: Undo function of a word processor

The issued commands are kept in a stack.

When you select "Undo", the **last** command is popped off the stack and undone.



■ Run-time stack that a processor or virtual machine:

Stores the values of variables in nested methods.

When a new method is called, its parameter variables and local variables are pushed onto a stack.

When the method exits, they are popped off again.

### Stack in the Java Library

 $\blacksquare$  Stack class provides push, pop and peek methods.

Table 7 Working with Stacks		
Stack <integer> s = new Stack&lt;&gt;();</integer>	Constructs an empty stack.	
s.push(1); s.push(2); s.push(3);	Adds to the top of the stack; s is now [1, 2, 3]. (Following the toString method of the Stack class, we show the top of the stack at the end.)	
<pre>int top = s.pop();</pre>	Removes the top of the stack; top is set to 3 and s is now [1, 2].	
head = s.peek();	Gets the top of the stack without removing it; head is set to 2. $ \\$	

63

### Queue

A queue

Lets you add items to one end of the queue (the tail)
Remove items from the other end of the queue (the head)
Items are removed in the same order in which they were added
First-in, first-out or FIFO order

■ To visualize a queue, think of people lining up.



■ Typical application: a print queue.

### Queue

- The Queue interface in the standard Java library has:
  - an  $\operatorname{\mathtt{add}}$  method to add an element to the tail of the queue,
  - a remove method to remove the head of the queue, and
  - a  ${\tt peek}$  method to get the head element of the queue without removing it.
- $_{\blacksquare}$  The <code>LinkedList</code> class implements the <code>Queue</code> interface.
- When you need a queue, initialize a Queue variable with a LinkedList object:

Quesectings q = new LinkedListO();
q.add(""); q.add(""); q.add("");
while (q.ide) > 0 ( system.out.print(q.temore)) + ""); ) // Prints A B C

Table 8 Working with Queues		
Queue <integer> q = new LinkedList&lt;&gt;();</integer>	The $\mbox{\tt LinkedList}$ class implements the Queue interface.	
<pre>q.add(1); q.add(2); q.add(3);</pre>	Adds to the tail of the queue; $q$ is now [1, 2, 3].	
<pre>int head = q.remove();</pre>	Removes the head of the queue; head is set to 1 and q is $[2, 3]$ .	
head = q.peek();	Gets the head of the queue without removing it; head is set to 2. $ \\$	

65

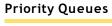
### **Priority Queues**

- A priority queue collects elements, each of which has a priority.
- Example: a collection of work requests, some of which may be more urgent than others.
- $_{\blacksquare}$  Does not maintain a first-in, first-out discipline.
- $_{\blacksquare}$  Elements are retrieved according to their priority.
- Priority 1 denotes the most urgent priority.

Each removal extracts the minimum element.

When you retrieve an item from a priority queue, you always get the most urgent one.

paul kline/iStockphot



 $\blacksquare$  Example: objects of a class  ${\tt WorkOrder}$  into a priority queue.

PriorityQueue<PortOrder> q = new PriorityQueue<P();
q.adidnew MorkOrder(3, "Bhampoo carpeta"));
q.addidnew MorkOrder(1, "Fix herbon sizk"));
q.addidnew MorkOrder(1, "Gordur cleaning supplies"));

- When calling q.remove() for the first time, the work order with priority 1 is removed.
- Elements should belong to a class that implements the Comparable interface.

Table 9 Working with Priority Queues	
<pre>PriorityQueue<integer> q =   new PriorityQueue&lt;&gt;&gt;();</integer></pre>	This priority queue holds Integer objects. In practice, you would use objects that describe tasks.
q.add(3); q.add(1); q.add(2);	Adds values to the priority queue.
<pre>int first = q.remove(); int second = q.remove();</pre>	Each call to remove removes the most urgent item: first is set to 1, second to 2.
<pre>int next = q.peek();</pre>	Gets the smallest value in the priority queue without removing it.

67

### **Self Check**

Why would you want to declare a variable as

instead of simply declaring it as a linked list?

Answer: This way, we can ensure that only queue operations can be invoked on the  $\, {\tt q} \,$  object.

Why wouldn't you want to use an array list for implementing a queue?

**Answer:** Depending on whether you consider the 0 position the head or the tail of the queue, you would either add or remove elements at that position. Both are inefficient operations because all other elements need to be moved.

69

# 

Why wouldn't you want to use a stack to manage print jobs?

**Answer:** Stacks use a "last-in, first-out" discipline. If you are the first one to submit a print job and lots of people add print jobs before the printer has a chance to deal with your job, they get their printouts first, and you have to wait until all other jobs are completed.