

Übung zur Vorlesung *Einführung in die Informatik 2 für Ingenieure (MSE)*

Alexander van Renen (renen@in.tum.de)

<http://db.in.tum.de/teaching/ss17/ei2/>

Lösungen zu Blatt 5

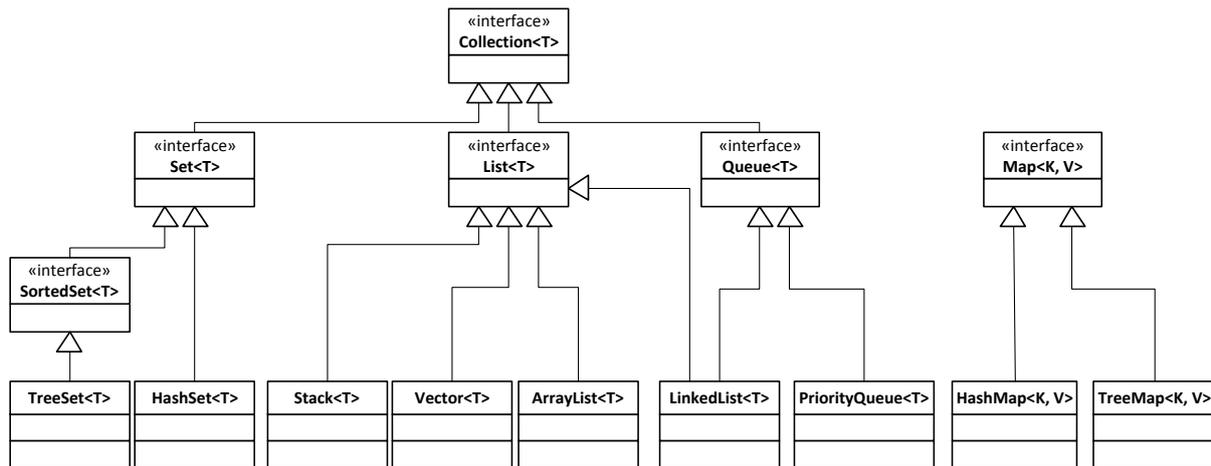


Abbildung 1: Ausschnitt aus dem Java Collections Framework (JCF)

Aufgabe 1: Java Collection Framework (JCF)

Das Java Collections Framework stellt viele häufig benötigte Datenstrukturen zur Verfügung, so dass Sie das Rad nicht neu erfinden müssen. Sie finden eine gute Einführung zu den Collections unter <http://docs.oracle.com/javase/tutorial/collections/>. In dieser Aufgabe geht es darum für verschiedene Anwendungsbeispiele jeweils die richtige Datenstruktur auszuwählen.

In dieser Aufgabe geht es um die Verwaltung von Filmen.¹ Auf der Webseite zur Vorlesung finden Sie vorbereitete Dateien. Sie sollen jeweils die effizienteste Datenstruktur wählen, wobei wir uns auf `PriorityQueue`, `HashMap`, `TreeMap` und `ArrayList` beschränken. Begründen Sie, welche Datenstruktur Sie gewählt haben und welche Nachteile die anderen gehabt hätten. Vervollständigen Sie den Java-Code von der Website entsprechend.

Tipp: Datenstrukturen wie die `PriorityQueue` und die `TreeMap` müssen ihre Elemente vergleichen können. Wenn also eigene Klassen in diesen Datenstrukturen verwendet werden (wie zum Beispiel `Movie`) dann muss angegeben werden wie Objekte dieser Klasse verglichen werden können. Dafür werden in Java sogenannte Comparatoren verwendet: <https://docs.oracle.com/javase/7/docs/api/java/util/Comparator.html>.

- a) Die zu implementierende Klasse `MovieRangeSearch` soll Filme nach ihrem Erscheinungsjahr verwalten. Die Methode `getMoviesBetweenYears(int fromYear, int toYear)` soll alle Filme zurückgeben, die in einem bestimmten Zeitraum erschienen sind.

¹Basierend auf der Liste der Top 250 Filme von [imdb.com](http://www.imdb.com/chart/top): <http://www.imdb.com/chart/top>

- b) Die Klasse `MovieTitleSearch` verwaltet Filme nach ihrem Namen. Dabei gibt die Methode `getMovieWithTitle(String title)` den Film mit dem angegebenen Titel zurück.
- c) Die Klasse `MovieVoteSearch` verwaltet Filme nach abgegebenen Bewertungen. Dabei soll die Methode `getMovieWithFewestVotes()` den Film mit den wenigsten Stimmen zurückgeben.
- d) Die Klasse `MovieRankSearch` verwaltet Filme anhand ihrer Platzierung in der Top 250 Liste von IMDB. Die Methode `getMovieForRank(int rank)` bestimmt für eine Position den Film.

Lösung 1

Die am besten passende Datenstruktur ist jeweils in **rot** gekennzeichnet.

- a) **TreeMap**. Die passende Datenstruktur für die `MovieRangeSearch` ist die `TreeMap`. Diese erlaubt effiziente Bereichszugriffe über die Methode `submap(from, to)`. Um eine Kollision bei Filmen die im gleichen Jahr produziert wurden zu vermeiden, sollte die `TreeMap` nicht auf einzelne Filme abbilden sondern auf eine Liste mit Filmen.

ArrayList, HashMap, PriorityQueue. Die anderen drei Datenstrukturen sind ungeeignet, da man bei ihnen für eine Bereichssuche jeden Film einzeln überprüfen muss. Da man jeden Film betrachten muss, benötigt dies bei n Filmen ungefähr n Operationen. In der Landau-Notation schreibt man dies als $\mathcal{O}(n)$.

- b) **HashMap**. Zur `MovieTitleSearch` passt am besten die Datenstruktur `HashMap`. Der Punktzugriff über die Methode `get()` benötigt nur konstante Zeit und ist somit unabhängig von der Anzahl der Filme. In Landau-Notation schreibt man dies als $\mathcal{O}(1)$.

TreeMap. Die `TreeMap` benötigt aufgrund der Baumeigenschaft bei n Filmen ungefähr $\log(n)$ Operationen für den Punktzugriff (Landau-Notation: $\mathcal{O}(\log(n))$). Der Grund dafür: Es muss von der Wurzel ein Pfad bis zum entsprechenden Blatt traversiert werden und der Baum hat eine Höhe von ungefähr $\log(n)$.

ArrayList, PriorityQueue. Bei der `ArrayList` und der `PriorityQueue` muss man über alle Elemente iterieren, um den passenden Film zu finden. Entsprechend ist die Laufzeit linear in der Anzahl der Filme (Landau-Notation: $\mathcal{O}(n)$).

- c) **PriorityQueue**. Die `PriorityQueue` verwaltet ihre Elemente nach Priorität. Dabei ist das Element mit der höchsten Priorität in der Wurzel verfügbar. Dies können wir ausnutzen, indem wir „weniger Stimmen“ als höhere Priorität definieren. Folglich ist somit in der Wurzel der `PriorityQueue` der Film mit den wenigsten Stimmen gespeichert und kann in konstanter Zeit abgefragt werden ($\mathcal{O}(1)$).

ArrayList, HashMap. Bei einer `HashMap` und einer `ArrayList` muss man über alle Filme iterieren, um manuell den mit den wenigsten Stimmen zu finden. Dies benötigt daher lineare Zeit ($\mathcal{O}(n)$).

TreeMap. Bei der `TreeMap` könnte man zumindest auf eine logarithmische Laufzeit hoffen, indem man die Anzahl an Stimmen als Schlüssel verwendet. Dies ist aber nicht möglich, da man keine zwei Elemente mit dem gleichen Schlüssel einfügen kann und es sehr wohl mehrere Filme mit der gleichen Anzahl Stimmen geben kann. Entsprechend müsste man ein anderes (eindeutiges) Attribut der Filmklasse als Schlüssel nehmen (z.B. den Titel wie bei Aufgabenteil a), das einem bei der Suche nicht hilft. Die Laufzeit ist dann wie bei `ArrayList` und `HashMap` linear in der Anzahl der Filme ($\mathcal{O}(n)$).

d) **ArrayList**. Die `ArrayList` ist hier am besten geeignet, da die Schlüssel dicht gepackt sind und sie die Operationen `get()` und `set()` mit konstanter Laufzeit ermöglicht ($\mathcal{O}(1)$). Im Gegensatz zu einem einfachen Array, erlaubt die `ArrayList` eine dynamische Größenveränderung. Man könnte Sie also problemlos vergrößern, wenn man statt den Top 250 die Top 500 Filme speichern wollte.

HashMap. Die `HashMap` bietet ebenfalls Punktzugriffe in $\mathcal{O}(1)$ und ist somit auch geeignet.

PriorityQueue. Die `PriorityQueue` erlaubt keinen Punktzugriff auf einen Schlüssel.

TreeMap. Die `TreeMap` benötigt logarithmische Zeit für einen Punktzugriff ($\mathcal{O}(\log(n))$).

Aufgabe 2: Generics

In dieser Aufgabe wollen wir einen generischen binären Suchbaum implementieren. Falls Sie sich noch unsicher mit Generics fühlen, können sie zunächst einen binären Suchbaum implementieren der lediglich Integer unterstützt und diesen dann in einem zweiten Schritt erweitern. Testen Sie Ihre Implementierung in jedem Fall mit geeigneten Beispielen.

1. Implementieren Sie einen binären Suchbaum,² der beliebige Elemente enthalten kann. Jeder Knoten sollte je einen Zeiger auf das linke und das rechte Kind haben. Der Baum sollte Methoden zum Einfügen und Suchen von Elementen anbieten. Wenn Sie eine Herausforderung suchen, können Sie auch noch Löschen implementieren.
2. Welches Problem gibt es, wenn man nacheinander die Zahlen 1, 2, 3, ..., 1.000.000 in aufsteigender Reihenfolge einfügt?

Lösung 2

Die folgende Klasse implementiert einen binären Suchbaum.

```
1 public class Binaerbaum<T extends Comparable<T>> {
2     // Der Schluessel
3     private T key;
4     // Das linke Kind
5     private BinaryTree<T> leftChild;
6     // Das rechte Kind
7     private BinaryTree<T> rightChild;
8     // Der Elternknoten
9     private BinaryTree<T> parent;
10
11     // Konstruktor
12     public BinaryTree(T key, BinaryTree<T> parent) {
13         this.key = key;
14         this.parent = parent;
15     }
16
17     // Fuege ein Element ein
18     public void insert(T element) {
19         insert(new BinaryTree<T>(element, null));
```

²Falls Ihnen entfallen ist, was ein Binärbaum ist: http://de.wikipedia.org/wiki/Binärer_Suchbaum

```

20 }
21
22 // Fuege einen Teilbaum an der richtigen Stelle ein
23 public void insert(BinaryTree<T> node) {
24     if (node.key.compareTo(key) <= 0) {
25         // Links einfüegen
26         if (leftChild == null) {
27             // Als linkes Kind einfüegen
28             leftChild = node;
29             node.parent = this; // Elternverweis korrekt setzen
30         } else {
31             // Beim linken Kind einfüegen
32             leftChild.insert(node);
33         }
34     } else {
35         // Rechts einfüegen
36         if (rightChild == null) {
37             // Als rechtes Kind einfüegen
38             rightChild = node;
39             node.parent = this; // Elternverweis korrekt setzen
40         } else {
41             // Beim rechten Kind einfüegen
42             rightChild.insert(node);
43         }
44     }
45 }
46
47 // Suche ein Element und gebe es aus, falls gefunden.
48 // Ansonsten gebe null zurueck.
49 public T search(T element) {
50     // Element gefunden?
51     if (element.compareTo(key) == 0) {
52         return element;
53     }
54     if (element.compareTo(key) < 0) {
55         // Links suchen
56         if (leftChild != null) {
57             return leftChild.search(element);
58         }
59     } else {
60         // Rechts suchen
61         if (rightChild != null) {
62             return rightChild.search(element);
63         }
64     }
65     // Element nicht gefunden
66     return null;
67 }

```

```

68
69 // Ersetze ein Kind durch einen Teilbaum
70 private void replaceChild(BinaryTree<T> child, BinaryTree<T>
    replacement) {
71     if (leftChild == child) {
72         leftChild = replacement;
73     } else if (rightChild == child) {
74         rightChild = replacement;
75     }
76 }
77
78 // Entferne den Knoten mit dem angegebenen Element
79 public void delete(T element) {
80     // Dieser Knoten muss geloescht werden
81     if (element.compareTo(key) == 0) {
82         // Fall 1: Keine Kinder, Knoten einfach loeschen
83         if (leftChild == null && rightChild == null) {
84             parent.replaceChild(this, null);
85         }
86         // Fall 2: Nur rechtes Kind, Knoten durch rechtes Kind ersetzen
87         else if (leftChild == null && rightChild != null) {
88             parent.replaceChild(this, rightChild);
89         }
90         // Fall 3: Nur linkes Kind, Knoten durch linkes Kind ersetzen
91         else if (leftChild != null && rightChild == null) {
92             parent.replaceChild(this, leftChild);
93         }
94         // Fall 4: Zwei Kinder
95         else {
96             // Knoten durch linkes Kind ersetzen
97             parent.replaceChild(this, leftChild);
98             // Rechtes Kind beim linken Kind einfuegen
99             leftChild.insert(rightChild);
100         }
101     }
102
103     // Ein Kindknoten muss geloescht werden
104     if (element.compareTo(key) <= 0) {
105         // Links loeschen
106         if (leftChild != null) {
107             leftChild.delete(element);
108         }
109     } else {
110         // Rechts loeschen
111         if (rightChild != null) {
112             rightChild.delete(element);
113         }
114     }

```

```
115 }
116 }
```

Aufgabe 3: For Each Loops

Das folgende Programm soll alle Primzahlen im Intervall $[0, 10.000]$ bestimmen.³ Leider wirft das Programm bei der Ausführung eine *Exception*. Wo und warum? Was ist eine mögliche Lösung?

```
1 import java.util.ArrayList;
2
3 public class Eratosthenes {
4
5     public static void main(String[] args) {
6         // Zahlen 2..10000 einfüegen
7         ArrayList<Integer> list = new ArrayList<Integer>();
8         for (int i = 2; i <= 10000; i++) {
9             list.add(i);
10        }
11        for (Integer zahl : list) {
12            // Probiere alle vorherigen Primzahlen als Teiler
13            for (Integer teiler : list) {
14                // Wenn der zu pruefende Teiler schon groesser als
15                // die Quadratwurzel ist , muss es eine Primzahl sein
16                if (teiler > Math.sqrt(zahl)) break;
17                // Zahlen mit Teiler sind keine Primzahlen und
18                // werden daher hier entfernt
19                if (zahl % teiler == 0) {
20                    list.remove(zahl);
21                    break;
22                }
23            }
24        }
25        // Alle verbliebenen Zahlen sind Primzahlen
26        for (Integer primzahl : list)
27            System.out.println(primzahl);
28    }
29 }
```

Lösung 3

In Zeile 21 wird ein Element aus der Liste gelöscht, über die wir in diesem Moment mit einer for-each-Schleife in Zeile 11 iterieren. Durch die Veränderung der Datenstruktur über die wir gleichzeitig iterieren, kann es zu Inkonsistenzen kommen, so dass beispielsweise ein gelöscht Element besucht würde. Erfreulicherweise gibt es im Java Collections Framework das Konzept des Iterator-Objekts mit dem wir das Problem leicht lösen können. Der Iterator erlaubt es ähnlich der for-each-Schleife eine Datenstruktur zu durchlaufen. Zusätzlich kann der Iterator durch sein

³Basierend auf dem Sieb des Eratosthenes: http://de.wikipedia.org/wiki/Sieb_des_Eratosthenes

Wissen über die Iteration, die er gerade verwaltet, eine `remove()`-Methode anbieten, die das zuletzt von `next()` zurück gegebenen Elements löscht – ohne dass es dabei zu Inkonsistenzen kommt. Entsprechend verwenden wir in der Lösung in Zeile 12 statt einer for-each-Schleife nun den Iterator und löschen in Zeile 24 die aktuelle Zahl ebenfalls über den Iterator aus der Liste.

```
1 import java.util.ArrayList;
2 import java.util.ListIterator;
3
4 public class EratosthenesLoesung {
5     public static void main(String[] args) {
6         // Zahlen 2..10000 einfüegen
7         ArrayList<Integer> list = new ArrayList<Integer>();
8         for (int i = 2; i <= 10000; i++) {
9             list.add(i);
10        }
11
12        ListIterator<Integer> iterator = list.listIterator();
13        while (iterator.hasNext()) {
14            Integer zahl = iterator.next();
15            // Probiere alle vorherigen Primzahlen als Teiler
16            for (Integer teiler : list) {
17                // Wenn der zu pruefende Teiler schon groesser als
18                // die Quadratwurzel ist, muss es eine Primzahl sein
19                if (teiler > Math.sqrt(zahl))
20                    break;
21                // Zahlen mit Teiler sind keine Primzahlen und
22                // werden daher hier entfernt
23                if (zahl % teiler == 0) {
24                    iterator.remove();
25                    break;
26                }
27            }
28        }
29
30        // Alle verbliebenen Zahlen sind Primzahlen
31        for (Integer primzahl : list)
32            System.out.println(primzahl);
33    }
34 }
```