BTI1-PTP, Aufgabenblatt 6

Sommersemester 2018

Sammlungen benutzen – Klassen als Objekte

Ausgabedatum: ..... 24. Mai 2018

Lernziele

Sammlungen (Interfaces und implementierende Klassen) des Java Collections Frameworks benutzen können; die Unterschiede zwischen einer Menge (`Set`) und einer Liste (`List`) kennen; Typtests und Typzusicherungen verstehen und programmieren können; die erweiterte for-Schleife für Sammlungen benutzen können; Java-Programme ohne BlueJ mit Hilfe der main-Methode von der Kommandozeile aufrufen können, Kommandozeilenparameter übergeben können.

Kernbegriffe

In praktisch allen Anwendungen werden *Sammlungen* gleichartiger Objekte manipuliert. Für die alltägliche Programmierung stellen Programmierbibliotheken meist Sammlungen als *dynamische Behälter* zur Verfügung, in die beliebig viele *Elemente* eingefügt und aus ihnen auch wieder entfernt werden können. Dabei sind zwei Eigenschaften für Klienten von solchen Sammlungen besonders relevant:

- Haben die Elemente in der Sammlung explizit eine *manipulierbare Reihenfolge* (wie bei einer Liste) oder ist ihre *Ordnung irrelevant* (wie beispielsweise bei einer allgemeinen Menge)?
- Sind *Duplikate* in der Sammlung zugelassen (wie bei einer Liste) oder darf ein Element nur einmal vorkommen (wie bei einer Menge)?

Wir konzentrieren uns hier zunächst auf diese Benutzungsaspekte von Sammlungen, indem wir Listen und Mengen benutzen.

Um für eine Sammlung entscheiden zu können, ob ein Element bereits enthalten ist, muss es ein Konzept von Gleichheit geben. Wir unterscheiden für Objekte *Gleichheit* von *Identität*: Zwei Objekte einer Klasse können *gleich* sein (etwa die gleichen Werte in ihren Exemplarvariablen haben), sind aber niemals *identisch* (ein Objekt ist nur mit sich selbst identisch). Gleichheit impliziert also nicht Identität, aber Identität impliziert Gleichheit: Wenn zwei Variablen/Referenzen auf *dasselbe* Objekt verweisen, verweisen sie automatisch auch auf *das gleiche* Objekt.

Alle Objekte in Java können auf Gleichheit miteinander verglichen werden, da an jedem Objekt die Operation `boolean equals(Object other)` aufgerufen werden kann. Sie ist in der Klasse `Object` definiert, die eine Handvoll Operationen definiert, die jedes Objekt in einem Java-System anbietet. Der Operation `equals` wird als Parameter das Exemplar mitgegeben, mit dem es verglichen werden soll. In Java-Klassen ist diese Methode standardmäßig als Prüfung auf Identität realisiert, sofern keine eigene `equals`-Methode implementiert wird.

Die *Sammlungsbibliothek* von Java (engl. *Java Collection Framework*, kurz JCF) stellt verschiedene Interfaces und Klassen für verschiedenartige Sammlungen zur Verfügung. Seit der Java-Version 5 ist es möglich, den Typ der Elemente einer Sammlung mit anzugeben. So deklariert beispielsweise

```
List<String> myList;
```

eine Variable `myList`, die nur auf Listen verweisen kann, die ausschließlich Strings enthalten, und `new ArrayList<String>()` erzeugt ein Exemplar der das Interface `List` implementierenden Klasse `ArrayList`, in dem nur Strings gespeichert werden können.

Ebenfalls seit der Java-Version 5 ermöglicht die *erweiterte for-Schleife* (engl.: for-each loop), die Elemente einer Sammlung zu durchlaufen. So gibt beispielsweise die folgende Schleife über die oben deklarierte Liste `myList` für jeden String in der Liste seine Länge aus:

```
for (String s : myList) // lies: für jeden String s in myList ...
{
    System.out.println(s.length());
}
```

Neben Set und List gibt es im JCF das Interface Map, das den Umgang mit *Abbildungen* modelliert. Eine Abbildung ist eine Sammlung von *Schlüssel-Wert-Paaren*, in der die Schlüssel eindeutig sein

müssen. Als Wert kann jeder Referenztyp dienen, beispielsweise auch eine Sammlung. Über die Operation `get` kann mit einem Schlüssel bequem auf einen gespeicherten Wert zugriffen werden.

Die Bibliotheken der Sprache Java sind in so genannten *Paketen* (engl.: packages) organisiert. Das Paket der Sammlungsbibliothek heißt `java.util`. Klassen, die Bibliotheksklassen und -Interfaces benutzen, importieren diese mit einer *Import-Anweisung* (z.B. `import java.util.ArrayList;`). Die Programmier-Schnittstelle (engl: *API – Application Programming Interface*) der verschiedenen Bibliotheken ist in der *API Specification* beschrieben, siehe http://docs.oracle.com/javase/8/docs/api/. Dort findet sich u.a. die Dokumentation der Sammlungsbibliothek und die der Operation `equals` aus der Klasse `Object` (im Paket `java.lang`).

In Java kann auch eine Klasse als ein Objekt angesehen werden, das zur Laufzeit einen Zustand haben und Operationen anbieten kann. Die Operationen eines Klassenobjektes (seine *Klassenoperationen*) werden mit dem Schlüsselwort `static` deklariert, ebenso wie mögliche Zustandsfelder für den Zustand des Klassenobjektes (seine *Klassenvariablen*). Die meisten Klassen in der Java-API sind jedoch zustandslos, so dass das Ergebnis einer Klassenoperation üblicherweise nur von ihren Parametern abhängt.

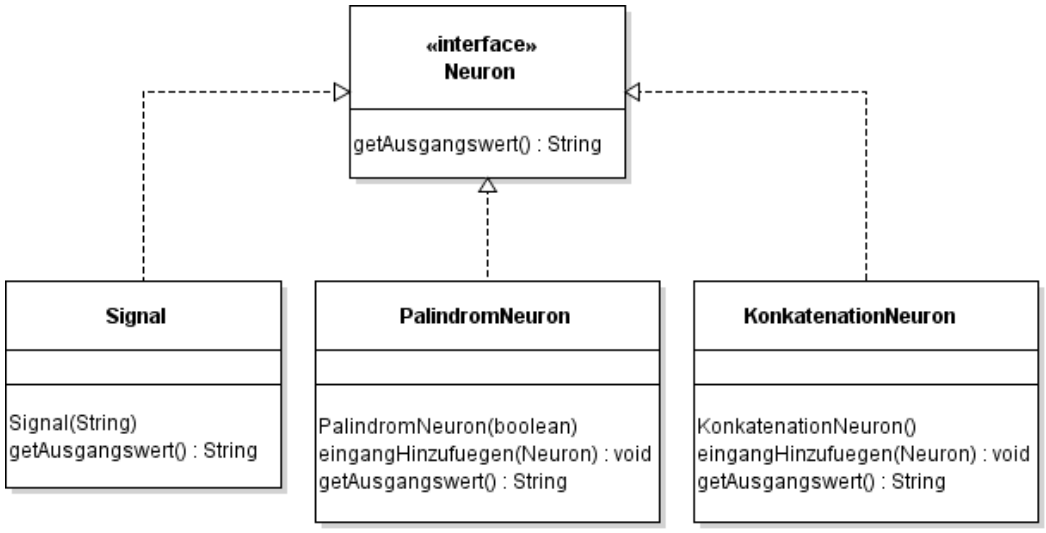
Arrays sind spezielle Sammlungen gleichartiger Elemente. Sie werden klassisch von imperativen Sprachen angeboten, meist unterstützt durch eine spezielle Syntax. Der Zugriff auf ein Element erfolgt, wie bei einer Liste, über einen Index. Da Arrays jedoch direkt auf den zugrunde liegenden Speicher abgebildet werden, kann mit den Mitteln der unterliegenden Rechnerarchitektur (mit Indexregistern o.ä.) ein sehr schneller wahlfreier Zugriff gewährleistet werden. Dafür sind Arrays jedoch in den meisten Sprachen statisch in ihrer Größe festgelegt, entweder bereits in ihrer Deklaration (wie in der Sprache Pascal) oder spätestens bei ihrer Erzeugung zur Laufzeit (wie in Java).

In Java können die Elemente eines Arrays sowohl Werte der Basistypen als auch Referenzen auf Objekte sein. Der Index ist in Java eine natürliche Zahl von 0 bis (Größe des Arrays)-1. Er wird in eckigen Klammern direkt hinter dem Bezeichner des Arrays verwendet (`a[0]` beispielsweise bezeichnet das erste Element des Arrays `a`).

Aufgabe 6.1 Neuronales Netz für Zeichenketten

In dieser Aufgabe entwickeln wir ein Netzwerk zur Verarbeitung von Zeichenketten, das durch den Aufbau *Neuronaler Netze* inspiriert ist. Im Kern eines solchen Netzes stehen *Neuronen*. Jedes Neuron hat mehrere Eingänge und einen Ausgang. Aus den Signalen an den Eingängen wird das Signal am Ausgang berechnet.

In unserer Anwendung werden Signale als Zeichenketten dargestellt. Außerdem drehen wir die Aktivierungsreihenfolge um: in der Realität sendet ein Neuron ein Signal am Ausgang, wenn ein Schwellwert von Eingangssignalen überschritten wird; wir hingegen „sammeln“ die Werte an den Eingängen und berechnen daraus den Ausgangswert, wenn dieser abgefragt wird. Das folgende Klassendiagramm gibt eine Übersicht, was insgesamt in 6.1 implementiert werden soll:



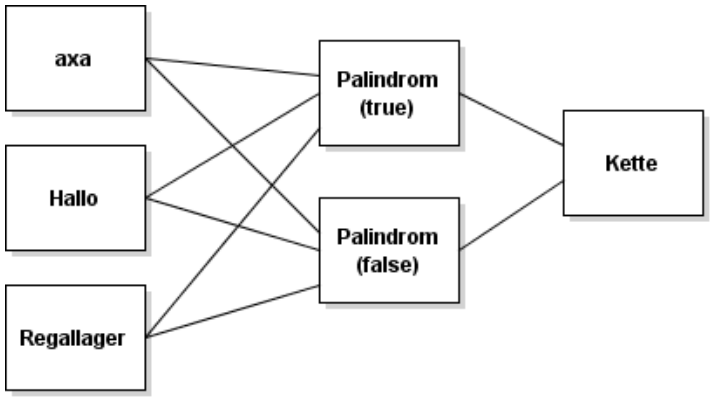
6.1.1 Wir benötigen zunächst ein Interface `Neuron` für alle Neuronen. Dieses bietet eine Operation `getAusgangswert` für den Zugriff auf den Wert am Ausgang (einen String).

Die Eingangswerte eines neuronalen Netzes werden durch spezielle Neuronen geliefert, die selber keine Eingänge haben: `Signal`. Diese liefern am Ausgang einen Wert, der über den Konstruktor gesetzt wird.
Alle anderen Neuronen bieten an ihrer Schnittstelle eine Operation `eingangHinzufuegen`, mit der jeweils ein Neuron als weiteres Eingangssignal hinzugefügt wird.

6.1.2 Ein `PalindromNeuron` ist in der Lage, Palindrome zu erkennen. Um seinen Ausgangswert zu bestimmen, prüft es seine Eingangswerte auf Palindrome. Ist nur einer der Eingangswerte ein Palindrom, wird dieser Wert zurückgegeben; liegen mehrere Palindrome an den Eingängen an, wird das längste geliefert. Ein `PalindromNeuron` kann in einem von zwei Modi arbeiten: Entweder wird Groß- und Kleinschreibung unterschieden oder nicht; diese Eigenschaft wird über einen Konstruktorparameter gesetzt.
Welcher Sammlungstyp ist hier für das Speichern der Eingänge geeignet?

6.1.3 Ein `KettenNeuron` verkettet die Werte seine Eingänge, in der Reihenfolge, in der sie hinzugefügt wurden. Zwischen je zwei Eingangswerten wird immer ein Leerzeichen eingefügt. Außerdem hat dieses Neuron ein Gedächtnis: es merkt sich seine vorherige Ausgabe und schickt diese zusätzlich seiner nächsten Ausgabe vorweg.
Welcher Sammlungstyp ist hier für das Speichern der Eingänge geeignet?

6.1.4 Als gute Softwareentwickler habt ihr natürlich(!) Testklassen zu euren Klassen geschrieben. Testet nun auch das Zusammenspiel der Klassen. Beispielsweise sollte das folgende Netzwerk…



… mit dem ersten Abfragen der Ausgabe des Ketten-Neurons liefern:

axa Regallager

und mit dem zweiten Abfragen:

axa Regallager axa Regallager

Testet diesen Fall und außerdem mindestens ein weiteres Netzwerk mit JUnit.

Aufgabe 6.2 Prägende Informatiker

6.2.1 Öffnet das Projekt *Informatiker* und studiert das Interface `Vergleicher`.

Erzeugt jeweils ein Exemplar der Klasse `PraegendeInformatiker` und der Klasse `PerNachname`. Ruft auf dem ersten Exemplar die Operation `schreibeGeordnet` auf und übergebt das zweite Exemplar als Parameter. Daraufhin sollten die Informatiker per Nachnamen geordnet auf der Konsole erscheinen.

Die Klasse `PerNachname` vergleicht offenbar anhand des Nachnamens. Wozu mag wohl die Klasse `PerAlter` gut sein? Probiert sie aus!

Möglicherweise ist euch aufgefallen, dass gleichaltrige Personen untereinander keine dem Benutzer sinnvoll erscheinende Reihenfolge haben. Wäre es nicht praktisch, wenn man mehrere Vergleicher miteinander kombinieren könnte? Also erst das Alter vergleichen, und bei gleichem Alter den Nachnamen?

Genau dafür gibt es die Klasse `Zweistufig`. Erstellt ein Exemplar und übergebt als Parameter jeweils ein Exemplar der Klassen `PerAlter` und `PerNachname`. Wenn ihr nun das Exemplar

der Klasse `Zweistufig` an `schreibeGeordnet` übergebt, sollte die Liste deutlich sinnvoller aussehen.

6.2.2 Schreibt eine Klasse `PerVorname`, welche die Vornamen der Personen miteinander vergleicht.

6.2.3 Schreibt eine Klasse `PerGeschlecht`, welche Frauen vor Männern einstuft. Schreibt auch eine JUnit-Testklasse dazu. Warum bietet sich dies hier an?

6.2.4 Könnt ihr durch geschickten Einsatz der Klasse `Zweistufig` auch 3 Vergleicher hintereinander schalten?

6.2.5 Schreibt eine Klasse `Umgekehrt`, welche sich genau umgekehrt zu einem anderen Vergleicher verhält. Beispielsweise soll `new Umgekehrt(new PerAlter())` junge Personen vor alten einstufen. Als Vorlage kann hierbei die (deutlich kompliziertere) Klasse `Zweistufig` dienen.

6.2.6 *Zusatzaufgabe*: Das Vergleichen zweier Objekte wie im Interface `Vergleicher` ist ein solch zentraler Gedanke, dass dafür in Java bereits das Interface `Comparator` im Paket `java.util` mitgeliefert wird:

https://docs.oracle.com/javase/8/docs/api/java/util/Comparator.html

Damit `Comparator` für beliebige Typen von Objekten funktionieren kann, gibt man in spitzen Klammern an, um welchen Typ es sich jeweils handelt. In unserem Fall würde man also `Comparator<Person>` schreiben. Dieses Parametrisieren von Typen mit anderen Typen wird erst im nächsten Semester ausführlich erläutert.

Legt eine Kopie des Projekts *Informatiker* an, indem ihr in BlueJ *Projekt / Speichern unter…* auswählt und dann einen beliebigen Namen wählt, z.B. *Informatiker2*.

Entfernt das Interface `Vergleicher` und verwendet stattdessen den Typ `Comparator<Person>`. Dazu müsst ihr in jedem Quelltext, der diesen Typ verwendet, folgendes in die erste Zeile schreiben:

```
import java.util.Comparator;
```

Aufgabe 6.3 Strings in der Kommandozeile analysieren (Termin 1)

In dieser Aufgabe wollen wir eine Java-Methode einmal nicht innerhalb von BlueJ aufrufen, sondern von der Kommandozeile des jeweiligen Betriebssystems. In Java ist für diesen Zweck eine spezielle Operation definiert worden: `public static void main(String[] args)`. Wenn eine Klasse eine Methode mit genau dieser Signatur definiert, dann kann diese Methode aus der Laufzeitumgebung der plattformabhängigen Java Virtual Machine aufgerufen werden.

6.3.1 Erstellt eine Klasse, die eine solche `main`-Methode enthält. Im Rumpf der Methode soll vorläufig lediglich eine beliebige Meldung mit `System.out.println` auf der Konsole ausgegeben werden. Ruft diese Methode in BlueJ auf.

6.3.2 Versucht nun, diese Methode von der Kommandozeile aus aufzurufen. Dazu müsst ihr zuerst ein Fenster öffnen, in dem ihr Kommandos eingeben könnt. (*Beispiel Windows: Start → Ausführen → cmd*) Wechselt in das Projekt-Verzeichnis von BlueJ, in dem eure Klasse (*Klassenname*.class) liegt. Dann könnt ihr folgende Zeile eingeben:

java <*Klassenname*>

6.3.3 In der Signatur der Methode `main` seht ihr, dass diese Parameter vom Typ `String` in Form eines String-Arrays entgegennimmt. Findet heraus, wie man aktuelle Parameter in der Konsole bei einem Aufruf der Methode übergeben kann. Ändert nun eure `main`-Methode so ab, dass die übergebenen Strings nacheinander mit `System.out.println` ausgegeben werden. Testet diese Änderung, indem ihr von der Kommandozeile aus eure `main`-Methode mit Parametern aufruft.

6.3.4 Schreibt in eurer Klasse eine Methode `analysiereText`, die für einen übergebenen String erfasst, wie häufig die 26 Buchstaben des Alphabets (ohne Umlaute, nur Kleinbuchstaben) darin vorkommen. Die Methode soll dazu auch ein `int`-Array der Länge 26 erhalten, das es entsprechend verändert:

```
void analysiereText(String text, int[] haeufigkeit)
```

Wendet eure Methode auf jeden Parameter der `main`-Methode an. Anschließend soll in der main-Methode das *Gesamtergebnis* für alle Parameter mit `System.out.println` ausgegeben werden. Testet erneut, entweder von der Kommandozeile aus oder in BlueJ. Tipp:

Für diese Aufgabe müssen Buchstaben auf Array-Positionen abgebildet werden. Dabei soll die Position 0 für den Buchstaben 'a' stehen, die Position 25 für den Buchstaben 'z'. Die korrekte Array-Position für einen Buchstaben erhaltet ihr, indem ihr 'a' subtrahiert.

6.3.5 *Zusatzaufgabe*: Was passiert, wenn ihr aus einer Klassenmethode (z.B. der `public static void main(String[] args`) auf eine Exemplarvariable zugreifen wollt? Gebt euren Betreuern eine Erklärung für das auftretende Verhalten.

Zusatzaufgabe 6.4 Das Geburtstagsparadoxon

Ein bekanntes mathematisches Rätsel, von dem ihr vielleicht schon einmal gehört habt, ist das *Geburtstagsparadoxon*. Dabei geht es um die Frage, wie wahrscheinlich es ist, dass in einer Gruppe von Personen mehrere Leute am gleichen Tag Geburtstag haben (wobei das Geburtsjahr keine Rolle spielt). Die Wahrscheinlichkeit ist schon für kleinere Gruppen wie etwa Partys und Schulklassen erstaunlich hoch. Das Geburtstagsparadoxon ist eine Veranschaulichung der allgemeinen Frage nach der Kollision von Zufallszahlen und spielt z.B. in der Kryptographie eine wichtige Rolle.

Es gibt mathematische Formeln, die die Wahrscheinlichkeit einer Kollision exakt berechnen. Diese werdet ihr in Veranstaltungen kennen lernen, die sich mit Kombinatorik und Stochastik beschäftigen. Als angehende Softwareentwickler wollen wir hier einen anderen Ansatz wählen. Wir simulieren eine große Anzahl von Partys mit Gästen und leiten aus den Messergebnissen einen empirischen Wert für die Wahrscheinlichkeit ab.

6.4.1 Öffnet das Projekt *Geburtstag* und schaut euch die *Dokumentation* der Klasse `Tag` an. Im Kommentar von `equals` steht, dass zwei `Tag`-Objekte, die den gleichen Tag darstellen, als gleich angesehen werden. `Tag`-Objekte, die nicht den gleichen Tag darstellen, sollen natürlich als ungleich angesehen werden.
Überprüft dies, indem ihr mindestens diese beiden Testfälle in der vorgegebenen JUnit-Testklasse implementiert! Wie weit lässt sich die Klasse `Tag` sinnvoll testen?

6.4.2 Schaut euch das Interface `Party` an. Hier werden zwei Operationen definiert: `fuegeGeburtstagHinzu` wird aufgerufen, wenn ein Gast seinen Geburtstag verraten hat.

`mindestensEinGeburtstagMehrfach` liefert `true`, sobald zwei Gäste am gleichen Tag Geburtstag haben.

Schreibt eine Klasse `PartyMenge`, die das Interface `Party` implementiert. Fügt dazu in der Methode `fuegeGeburtstagHinzu` den übergebenen Geburtstag in ein `HashSet` von Geburtstagen ein. Das Interface `Set` definiert die Schnittstelle einer Menge. Eine Menge enthält keine Duplikate und die Elemente in einer Menge haben keine explizite Reihenfolge (bzw. die gekapselte, interne Reihenfolge ist nicht relevant für den Umgang mit einer Menge). Wie bemerkt ihr, ob ein Geburtstag bereits enthalten ist?

http://docs.oracle.com/javase/7/docs/api/java/util/Set.html

Testet eure Implementation, indem ihr ein Exemplar von `Simulation` erstellt und daran die Methode `test()` aufruft. Falls das Ergebnis `false` ist, habt ihr einen Fehler gemacht.

6.4.3 Schätzt zunächst, wie viele Gäste ungefähr nötig sein müssten, um Kollisionswahrscheinlichkeiten von 50%, 75% und 95% zu bekommen. Testet eure Vermutungen anschließend mit der Methode `simuliere(int gaeste)`.

Überlegt, wann die Wahrscheinlichkeit einer Kollision *exakt* 100% sein muss.

6.4.4 Kommentiert die Methode `equals` in der Klasse `Tag` aus (oder benennt sie einfach um). Führt nun eine Simulation mit 366 Gästen aus. Was beobachtet ihr? Woran liegt das? Schaut euch ggf. die Dokumentation der Methode `equals` in der Klasse `Object` an. **Sorgt anschließend dafür, dass `equals` wieder korrekt funktioniert** (Auskommentierung bzw. Umbenennung rückgängig machen).

Danksagung

Die Aufgaben wurden freundlicher Weise von Prof. Dr. Schmolitzky aus früheren Veranstaltungen zur Verfügung gestellt.