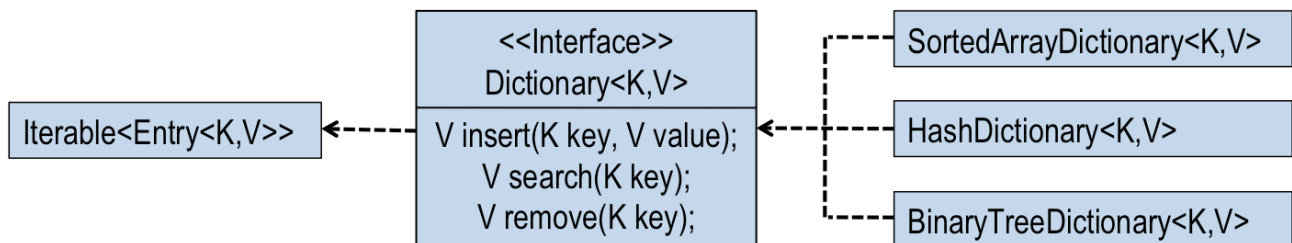


Aufgabenblatt 1



- Das Interface **Dictionary** (siehe Web-Seite) definiert Methoden zur Verwaltung von Datensätzen (Entries), die aus einem Schlüssel vom Typ K und Nutzdaten vom Typ V bestehen. Außerdem können über die Datensätze eines Dictionary iteriert werden (Iterable). Beachten Sie, dass Dictionary bereits eine Klasse Entry zur Verfügung stellt, die auch bei eigenen Klassenimplementierungen verwendet werden kann.
- **SortedArrayDictionary** implementiert ein Dictionary mit einem Feld, in dem die Datensätze lückenlos und sortiert gespeichert werden. Für die Suche wird binäre Suche eingesetzt.
- **HashDictionary** verwendet als Implementierung eine wie in der Vorlesung besprochene Hashtabelle mit linear verketteten Listen. Achten Sie darauf, dass die Größe der Hashtabelle eine Primzahl ist. Wird ein bestimmter Füllungsgrad (load factor) z.B. von 2 überschritten, dann wird die Tabelle vergrößert, so dass die neue Größe etwa doppelt so groß und wieder eine Primzahl ist. Die Daten werden dann sofort umkopiert.
- **BinaryTreeDictionary** setzt für den Iterator die in der Vorlesung besprochenen AVL-Bäume mit einer Eltern-Zeiger-Technik ein. Zu Testzwecken soll eine prettyPrint-Methode verwendet werden (siehe rudimentäre Klasse BinaryTreeDictionary auf der Web-Seite).
Hinweis:
 Erweitern Sie zuerst die binären Suchbäume um die Elternzeigertechnik. Prüfen Sie mit Hilfe von prettyPrint die Korrektheit der Elternzeiger. Testen Sie auch die Korrektheit des Iterators. Erweitern Sie nun die binären Suchbäume zu AVL-Bäumen, indem Rotationsoperationen durchgeführt werden. Beachten Sie dabei, dass auch bei den Rotationsoperationen die Elternzeiger entsprechend gesetzt werden müssen.

Es sind folgende Aufgabenteile zu lösen:

1. Implementieren Sie alle Klassen in Java und testen Sie die Klassen ausgiebig. Verwenden Sie dazu das zur Verfügung gestellte Testprogramm (siehe Web-Seite).
2. Schreiben Sie eine textbasierte Benutzerschnittstelle für eine Wörterbuch-Anwendung Deutsch-Englisch mit folgender Funktionalität:

Konsolen-Kommando	Bedeutung
<code>create</code> <i>Implementierung</i>	Legt ein Dictionary an. SortedArrayDictionary ist voreingestellt.
<code>r [n] Dateiname</code>	Liest (read) die ersten n Einträge der Datei in das Dictionary ein. Wird n weggelassen, dann werden alle Einträge eingelesen. Einfachheitshalber kann ein JFileChooser-Dialog verwendet werden (siehe Prog2, GUI). Dann wird aber der Dateiname im Kommando weggelassen.
<code>p</code>	Gibt alle Einträge des Dictionary in der Konsole aus (print).
<code>s deutsch</code>	Gibt das entsprechende englische Wort aus (search).
<code>i deutsch englisch</code>	Fügt ein neues Wortpaar in das Dictionary ein (insert).
<code>d deutsch</code>	Löscht einen Eintrag (delete).
<code>exit</code>	beendet das Programm.

3. Die verschiedenen Implementierungen sollen einer Performance-Untersuchung unterzogen werden. Messen Sie dazu die CPU-Zeiten für verschiedene Anwendungsfälle für die gegebene Wörterbuchdatei mit knapp $n = 16000$ Einträgen und tragen Sie die Zeiten in folgende Tabelle ein. ms

	SortedArray -Dictionary	Hash- Dictionary	BinaryTree- Dictionary
Aufbau eines Wörterbuchs mit $n = 8000$ Einträgen	132032.215000	6177.892000	7573.576000
Aufbau eines Wörterbuchs mit $n \approx 16000$ Einträgen	498869.280000	12903.743000	25351.651000
Erfolgreiche Suche ¹⁾ für $n = 8000$	15.097000	54.081000	16.266000
Erfolgreiche Suche ¹⁾ für $n \approx 16000$	13.125000	50.775000	14.576000
Nicht erfolgreiche Suche ²⁾ für $n = 8000$	7.275000	7.466000	5.119000
Nicht erfolgreiche Suche ²⁾ für $n \approx 16000$	5.939000	4.132000	3.844000

¹⁾ Bei der erfolgreichen Suche wird jedes deutsche Wort des Wörterbuchs genau einmal gesucht und die Gesamtzeit gemessen. Testen Sie mit einem Wörterbuch mit $n = 8000$ und dann mit $n \approx 16000$ Einträgen. Schreiben Sie dazu vor der Laufzeitmessung alle deutschen Wörter in eine Liste und iterieren dann bei der Laufzeitmessung über diese Liste.

²⁾ Bei der nicht erfolgreichen Suche kann einfach nach jedem englischen Wort gesucht werden, das ja als deutsches Wort so gut wie nicht vorkommt. Auch hier ist die Gesamtzeit für $n = 8000$ bzw. $n \approx 16000$ einzutragen. Schreiben Sie dazu vor der Laufzeitmessung alle englischen Wörter in eine Liste und iterieren dann bei der Laufzeitmessung über diese Liste.