

**Hinweise zur Bearbeitung der Klausur zum Kurs 1661 Datenstrukturen I**

Bitte lesen Sie sich diese Hinweise vollständig und aufmerksam durch, bevor Sie mit der Bearbeitung der Klausur beginnen. Beachten Sie insbesondere die Punkte 7 und 8!

1. Die Klausurdauer beträgt 2 Stunden.
2. Prüfen Sie bitte die Vollständigkeit Ihrer Unterlagen. Die Klausur umfasst:
  - 2 Deckblätter
  - diese Hinweise
  - 1 Formblatt für eine Teilnahmebescheinigung zur Vorlage beim Finanzamt
  - 7 Aufgaben auf den Seiten 2–5
3. Bevor Sie mit der Bearbeitung der Klausuraufgaben beginnen, füllen Sie bitte die folgenden Teile der Klausur aus:
  - (a) sämtliche Deckblätter mit Name, Anschrift sowie Matrikelnummer. Markieren Sie vor der Abgabe auf allen Deckblättern die von Ihnen bearbeiteten Aufgaben.
  - (b) die Teilnahmebescheinigung, falls Sie diese wünschen.
4. Schreiben Sie Ihre Lösungen auf Ihr eigenes Papier (DIN A4) und nicht auf die Seiten mit den Aufgabenstellungen. Heften Sie vor Abgabe der Klausur die Deckblätter und die Aufgabenstellungen (und evtl. die Teilnahmebescheinigung) vor Ihre Bearbeitung. **Sie müssen die Aufgabenstellung also mit abgeben!** Sie erhalten die Aufgabenstellungen zusammen mit Ihrer korrigierten Klausur zurück.
5. Schreiben Sie bitte auf jedes Blatt Ihrer Bearbeitung oben links Ihren Namen und oben rechts Ihre Matrikelnummer. Nummerieren Sie Ihre Seiten bitte durch.
6. Wenden Sie bei der Lösung der Aufgaben, soweit dies möglich ist, die Algorithmen und Notationen aus den Kurseinheiten an.
7. Schreiben Sie, wenn Algorithmen gefordert sind, keine kompletten Java-Programme, sondern beschränken Sie sich auf die wesentlichen Teile des Algorithmus, die z.T. auch in Prosa formuliert werden können. Formulieren Sie Algorithmen aber so, dass die elementaren Einzelschritte erkennbar werden.  
  
Sparen Sie bei solchen Aufgaben nicht mit Kommentaren. Wenn Ihre Lösung aufgrund fehlender Kommentare nicht verständlich ist, führt das zu Punktabzug.
8. Vermeiden Sie in jedem Fall bei der Definition von Funktionen in Algebren Java-Programme sowie Algorithmen. Geben Sie lediglich mathematische Definitionen an, wie sie im Kurstext verwandt worden sind!  
  
Ebenso sind Mengendefinitionen in mathematischer Mengennotation durchzuführen. Geben Sie auf keinen Fall Datentypdefinitionen an und verwenden Sie keine konkreten Datenstrukturen!
9. Als Hilfsmittel sind nur unbeschriebenes Konzeptpapier und Schreibzeug zugelassen. Die Reinschrift der Klausur darf **nicht mit Bleistift** erfolgen.
10. Durch Lösen der Aufgaben sind maximal 100 Punkte erreichbar. Zusätzlich werden Ihnen ggf. bis zu 22 Bonuspunkte für korrektes Lösen der Einsendeaufgaben angerechnet. Sie dürfen damit rechnen einen Übungsschein bzw. ein Zertifikat zu erhalten, wenn Sie insgesamt mindestens 60 Punkte erreichen.

**30 Punkte Aufgabe 1 Algebra**

In dieser Aufgabe sollen Sie eine Algebra angeben, mit der es möglich ist, eine DVD-Sammlung zu beschreiben und Operationen auf dieser auszuführen.

Eine DVD besteht aus einer Kennung (nennen wir sie EAN) und enthält Filme. Die EAN ist eine natürliche Zahl. Vereinfachend gehen wir davon aus, dass die EAN innerhalb einer Sammlung eindeutig ist, d.h. keine doppelten DVDs vorhanden sind. Allerdings kann der gleiche Film auf unterschiedlichen DVDs existieren.

Für einen Film sollen folgende Daten hinterlegt werden:

der Name des Films,  
der Regisseur/die Regisseurin (ggf. mehrere oder keiner),  
die Laufzeit in ganzen Sekunden

Als Operationen der Algebra sollen (mindestens) bereitgestellt werden:

1. *createMovie* und *createDVD*: erzeugt aus den notwendigen Daten das entsprechende Objekt,
2. *createCollection* erzeugt eine leere Sammlung,
3. *getName*, *getDirector*, *getRuntime*: dienen zur Abfrage der gewünschten Information zu einem Film,
4. *insertDVD*: fügt eine DVD in die Sammlung ein,
5. *containsMovie*: prüft für eine Sammlung, ob ein – nur durch seinen Namen dargestellter – Film in einer Sammlung enthalten ist,
6. *singleMovies*: wählt aus einer Sammlung diejenigen Filme aus, die sich allein auf einer DVD befinden, d.h. die DVD enthält nur diesen einen Film,
7. *longDVDs*: ermittelt aus einer Sammlung diejenigen DVDs, deren Gesamtspielzeit größer ist, als ein gegebener Wert,
8. *moviesOf*: ermittelt in einer DVD-Sammlung alle Filme, bei denen ein gegebener Regisseur/eine gegebene Regisseurin beteiligt war,
9. *avgRuntimeOf*: ermittelt die durchschnittliche Laufzeit aller Filme eines bestimmten Regisseurs in einer Sammlung.

Zur Vereinfachung dieser Aufgabe können Sie davon ausgehen, dass die Datentypen **int**, **real**, **bool**, **string** mit den üblichen Operatoren sowie der parametrisierbare Datentyp **set** mit den Operationen *empty*, *insert*, *delete*, *member*, *size* bekannt sind und verwendet werden können. Damit kann z.B. die Trägermenge *string* des Typs **string** verwendet werden.

- 5 Punkte (a) Geben Sie die Sorten und die Operatorsignaturen der Algebra an.
- 10 Punkte (b) Geben Sie die Trägermengen für die Sorten der Algebra an.
- 15 Punkte (c) Geben Sie die Funktionen für alle Operatoren der Algebra an. Sie können auf die Punktnotation zum Zugriff auf die Elemente von Tupeln zurückgreifen.

**Aufgabe 2      Dictionaries****8 Punkte**

**Achtung:** Bei dieser Aufgabe erhalten Sie für jede richtige Antwort Punkte angerechnet, allerdings werden Ihnen ebenso für jede falsche Antwort Punkte **abgezogen**. Insgesamt können Sie bei dieser Aufgabe jedoch minimal 0 Punkte, maximal 8 Punkte erhalten.

- (a) Welche Operationen stellt der Datentyp *Dictionary* zur Verfügung und was stellt er dar? 1 Punkt
- (b) Nennen Sie 7 Implementierungsmöglichkeiten aus dem Kurstext für den Datentyp *Dictionary*. Geben Sie jeweils die worst-case Laufzeit für die typischen *Dictionary*-Operationen für jede Implementierung an. Begründungen der Laufzeiten sind nicht erforderlich. 7 Punkte

**Aufgabe 3      Hashing****10 Punkte**

Fügen Sie die Namen

*Karl, Barbara, Boris, Anton, Karla, Katrin, Peter, Diana*

in dieser Reihenfolge in eine Hashtabelle mit  $m=10$  ein. Benutzen Sie dabei die Divisionsmethode als Hashfunktion und die Grundform des quadratischen Sondierens (nur Addition). Den Namen wird zur Berechnung der Hashfunktion als Wert jeweils die Summe der Buchstabenwerte ihrer ersten 3 Zeichen zugeordnet, wobei „A“ mit 1, „B“ mit 2, ..., „Z“ mit 26 bewertet ist. Geben Sie für jeden eingefügten Namen die Folge der Behälternummern an, die getestet wird. Wie lautet der Inhalt der finalen Hashtabelle?

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

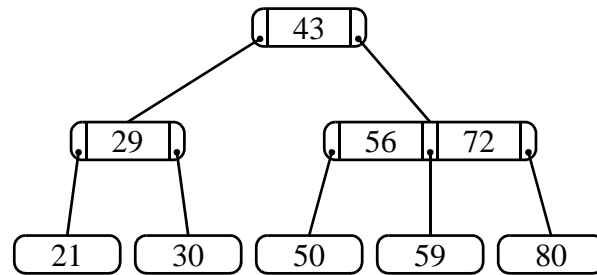
**Aufgabe 4      HeapSort****15 Punkte**

Sortieren Sie die unten stehende Folge absteigend mittels Heapsort. Geben Sie zunächst den initialen Heap an. Geben Sie weiterhin den Heap nach jedem Einsinkenlassen eines Elements an. Zusätzlich zum Heap ist auch noch die bisher sortierte Teilfolge anzugeben. Die Angabe des Heaps kann wahlweise graphisch oder als Array-Einbettung erfolgen.

35, 68, 75, 42, 56, 42, 31, 12, 19, 87, 77

### 15 Punkte Aufgabe 5 B-Bäume

- 1 Punkt (a) Von welcher Ordnung ist der folgende B-Baum?



- 7 Punkte (b) Löschen Sie aus dem Baum aus Teilaufgabe (a) folgende Schlüssel in der gegebenen Reihenfolge:

80, 43, 21.

Zeichnen Sie den Baum vor und nach jeder Ausgleichoperation und markieren Sie den Knoten, bei dem eine B-Baum-Eigenschaften verletzt wird.

- 7 Punkte (c) Fügen Sie folgende Schlüssel in der gegebenen Reihenfolge in einen leeren B-Baum der Ordnung 1 ein:

21, 59, 30, 29, 56, 80, 50, 43, 72.

Zeichnen Sie den Baum jeweils vor und nach einer Split-Operation sowie den vollständig gefüllten Baum. Markieren Sie Knoten, bei denen B-Baum-Eigenschaften verletzt werden. Sind beide Nachbarn bei einer Balance-Operation geeignet, so soll der linke Nachbar verwendet werden.

**Aufgabe 6      Intersection für B-Bäume****21 Punkte**

In dieser Aufgabe geht es darum, einen effizienten Algorithmus für folgende Operation auf zwei B-Bäumen der Ordnung  $b$  mit Schlüsseltyp  $T$  zu entwickeln:

$$\begin{aligned} \textit{intersection}: BTree(T) \times BTree(T) &\rightarrow \textit{set}(T) \\ \textit{intersection}(B_1, B_2) &= \{k \mid \textit{member}(B_1, k) \wedge \textit{member}(B_2, k)\} \end{aligned}$$

Es sei  $n$  die Anzahl der Schlüssel im ersten,  $m$  die der Schlüssel im zweiten B-Baum.

- (a) Welches Verfahren wurde im Kurstext vorzugsweise verwendet, um die Operation *intersection* auf geordneten Darstellungen von Mengen schnellstmöglich auszuführen und dabei eine geordnete Ausgabemenge zu erhalten? Beschreiben Sie das Verfahren kurz und informell. Was ist dabei die beste erzielbare Zeitkomplexität im worst-case? Wie kann man diese gegebenenfalls unter Berücksichtigung der Datenorganisation der Mengendarstellung (z.B. für B-Bäume) verbessern? Begründen Sie jeweils! 5 Punkte
- (b) Welche im Kurstext vorgestellte Hilfsstruktur können Sie verwenden, um einen In-Order-Durchlauf auf einem B-Baum der Ordnung  $b$  mit  $n$  Schlüsseln in  $O(n)$  Zeit ohne rekursive Aufrufe zu realisieren? Denken Sie dabei an das im Kurstext vorgestellte Verfahren zur Eliminierung rekursiver Aufrufe. Wie ist die Speicherplatzkomplexität für diese Datenstruktur in diesem Anwendungsfall? 1 Punkt
- (c) Durchläufe durch Datenstrukturen realisiert man oft mittels besonderer Hilfs-Datenstrukturen, sogenannter *Iteratoren*. Diese haben zwei Operationen: *init(B)*, um einen Durchlauf durch die Datenstruktur  $B$  vorzubereiten und benötigte interne Datenstrukturen zu initialisieren, und *next()*, um das nächste Element aus  $B$  oder, falls es keines mehr gibt, NULL zurückzuliefern. Beschreiben Sie effiziente Algorithmen für *init* und *next* eines Iterators für B-Bäume der Ordnung  $b$ . Verwenden Sie dabei die Hilfsdatenstruktur aus Teilaufgabe (b). 11 Punkte
- (d) Nutzen Sie die Erkenntnisse aus (a)–(c) und skizzieren Sie einen Algorithmus, der die oben beschriebene Operation *intersection* in  $O(n + m)$  Laufzeit realisiert und höchstens  $O(\log n + \log m)$  Platz für Hilfsstrukturen verwendet. Begründen Sie Laufzeit und Platzbedarf. 4 Punkte

**Aufgabe 7      Deckblatt****1 Punkt**

Lesen Sie sich die „Hinweise zur Bearbeitung“ sorgfältig durch. Füllen Sie den dort aufgeführten Anweisungen entsprechend beide Deckblätter vollständig und korrekt aus.