

29. Bundeswettbewerb Informatik 2010/2011 2. Runde

Tim Taubner, Verwaltungsnummer 29.108.01

21. März 2011

Dies ist die Dokumentation zu den von mir bearbeiteten Aufgaben 1 und 2 der 2. Runde des 29. Bundeswettbewerbs Informatik 2010/2011. Die mir zugeteilte Verwaltungsnummer ist 29.0108.01. Für alle Aufgabe werden jeweils die Lösungsidee und eine Programm-Dokumentation angegeben, sowie geeignete Programm-Ablaufprotokolle und der Programm-Text selbst. Auf die ausführbaren Lösungen wird in der Dokumentation verwiesen. Der Quelltext ist beigelegt. Ebenfalls enthalten sind weiterführende Gedankengänge, diese erhalten ebenfalls einen eigenen Unterpunkt. In diesem ist sowohl kurz die Idee als auch die Implementation erläutert. Zusätzlich ist am Ende eine allgemeine Beschreibung enthalten, wie die erstellten Programme von der mitgelieferten CD aus gestartet werden können. Alle eingereichten Quelldateien, Kunsterzeugnisse (wie z.B. Bilder) und ausführbare Programmdistributionen wurden alleine von mir, Tim Taubner, erstellt.

Inhaltsverzeichnis

A. Allgemeines	2
1. Persönliche Anmerkungen	2
2. Dateistruktur der CD	2
3. Ausführvoraussetzungen	3
4. Starten der Programme	3
B. Aufgabe 1: Kisten in Kisten in Kisten	4
1. Lösungsidee	4
1.1. Allgemein	4
1.2. Bruteforce	4
1.3. Bruteforce nach Aufteilung	5
1.4. Online Packer	5
2. Implementierung	6
3. Programmabläufe	7
3.1. Algorithm-Contest (Packdichte)	7
3.2. Algorithm-Contest (Laufzeit)	7
4. Programmtext	7
5. Programmnutzung	7
C. Zweite bearbeitete Aufgabe: (2) Containerklamüsel	8
1. Lösungsidee	8
1.1. Vorüberlegungen	8
1.2. Algorithmen	8
2. Implementation	8

A. Allgemeines

1. Persönliche Anmerkungen

Der BWInf und ich Der Bundeswettbewerb Informatik konnte mich sehr begeistern. Viel konnte ich bereits durch die 1. Runde lernen, z.B. wie eine gute Dokumentation erstellt werden kann. Auch das Ergebnis lässt sich sehen. Wenn ich gefragt werde was ich eigentlich am PC mache, klappe ich mein Laptop auf und zeige die Dokumentation zur 1. Aufgabe¹. Viel besser kann man finde ich nicht zeigen, dass Informatik *nicht* nur Programmieren ist.

Wahl der Programmiersprache Die benutzte Programmiersprache ist durchgehend Scala. Das liegt einfach an meiner Neigung, kurzen und dichtgepackten² Code zu schreiben. Ich bin mir durchaus bewusst, dass Scala - noch - keine weit verbreitete Sprache ist. Aber ich denke, dass es auch einem Scala-fremden Informatiker gefällt, wenn aussagekräftiger Code abgegeben wird.

Dank Ich erlaube mir hier, Personen zu danken, die mir zu dieser Einsendung verholfen haben. Auch wenn alle Leistungen im Sinne des Wettbewerbs von mir erbracht wurden, habe ich dazu nicht wenig Energie aus der Umgebung gezogen.³ Zum einen meiner Freundin⁴, aber auch meiner Familie. Sie scheinen bereits ein Algorithmus entwickelt haben, mit meinen einsilbigen Antworten fertig zuwerden.

2. Dateistruktur der CD

Die Dateien auf der CD sind folgendermaßen strukturiert. Jede Aufgabe hat einen Ordner AufgabeX mit den beiden folgenden Unterordnern.

¹siehe Einsendung zur 1. Runde, Einsendungsnr. 108, besonders Aufgabe 1

² Hier zeigt sich eine Schwäche der deutschen Sprache: Sie ist *verbose*. Ich versuche hier *concise* aus dem Englischem zu übersetzen

³ Vergleichen Sie dies mit einem Eisberg, er zieht beim schmelzen die ganze Wärme aus der Umgebung.

⁴Meine eigene Perle der Informatik ;-]

src Unterordner, in dem die Quelltexte in der Paketstruktur (de/voodle/..) liegen

dist Unterordner, in dem ausführbare Dateien oder - wie bei Aufgabe 1 - andere Erzeugnisse sowie benötigte Bibliotheken enthalten sind

Zusätzlich ist die vorliegende Dokumentation digital unter **TeX-Doku-Einsendung-2910801-Tim-Taubner.pdf** im Wurzelverzeichnis zu finden. Auch die L^AT_EX-Quelldateien, mit der diese Dokumentation erzeugt wurden, ist im Verzeichnis "TeX-Doku-Quelldateien" zu finden.

3. Ausführvoraussetzungen

Um die Programmbeispiele ausführen zu können, müssen folgende Voraussetzungen erfüllt werden:

Java Runtime: Mindestens Version 1.5 (Java 5), empfohlen: ≥ 1.6

Prozessor: Mindestens 1 GHz, empfohlen: ≥ 1.6 GHz

Arbeitsspeicher: Mindestens 512 MB, empfohlen: ≥ 1 GB

Grafikkarte: beliebig

Getestete Betriebssysteme: Windows 7, Windows XP, Linux (Ubuntu 10.04, Kubuntu 10.10)

4. Starten der Programme

Wurzelverzeichnis Im Wurzelverzeichnis der CD beginnt die Ordnerhierarchie der mitgelieferten Dateien. Stellen Sie bitte sicher, dass Sie im Wurzelverzeichnis sind, bevor Sie die in den jeweiligen Aufgaben beschriebene Startanleitungen ausführen. (z.B. durch neues Starten der Kommandozeile gemäß folgender Anleitung)

Starten der Kommandozeile Da die meisten mitgelieferten Programme aus der Kommandozeile gestartet werden müssen, soll hier kurz erläutert werden, wie Sie die Kommandozeile unter den gängigeren Betriebssystemen starten können.⁵

Unter Windows Unter Windows starten Sie die Kommandozeile durch: Start → Ausführen → 'cmd' eingeben → Kommandozeile. Nun können Sie durch Angabe des Laufwerksbuchstabens des CD-Laufwerks (z.B. "E:") auf das Wurzelverzeichnis der CD wechseln. Alle Befehle können nun durch Copy&Paste entsprechend der Nutzungsdokumentation der jeweiligen Aufgabe ausgeführt werden. Z.B. für Aufgabe 3 mit:

```
java -jar Aufgabe1/target/Kisten.jar
```

Unter GNOME Unter gängigeren GNOME Distributionen wie z.B. Ubuntu 8 starten sie die Kommandozeile durch: Applikationen → System → Terminal. Die CD wird unter Standard-distributionen unter /media/disk o.ä. eingehängt. Wechseln Sie durch 'cd /media/disk' in das Wurzelverzeichnis der CD. Alle Befehle können nun durch Copy&Paste entsprechend der Nutzungsdokumentation der jeweiligen Aufgabe ausgeführt werden. Z.B. für Aufgabe 3 mit:

```
java -jar Aufgabe1/target/Kisten.jar
```

Unter KDE Unter gängigeren KDE-Distributionen wie z.B. Kubuntu 9 starten sie die Kommandozeile durch: Start → Applikationen → System → Terminal. Die CD wird unter Standard-distributionen unter /media/disk o.ä. eingehängt. Wechseln Sie durch 'cd /media/disk' in das Wurzelverzeichnis der CD. Alle Befehle können nun durch Copy&Paste entsprechend der Nutzungsdokumentation der jeweiligen Aufgabe ausgeführt werden. Z.B. für Aufgabe 3 mit:

```
java -jar Aufgabe1/target/Kisten.jar
```

Unter Mac OS X Leider steht mir kein Mac zur Benutzung bereit, das Öffnen der Konsole sollte jedoch entweder selbsterklärend oder ähnlich der unter KDE/GNOME sein. (Beachten Sie bitte, dass Aufgabe2 leider nicht unter Mac OS X lauffähig ist)

⁵Ich respektiere Ihre wahrscheinlich umfassenden Kenntnisse mit Perl. Die Anleitung zum Kommandozeilestart dient lediglich der Vollständigkeit.

B. Aufgabe 1: Kisten in Kisten in Kisten

1. Lösungsidee

1.1. Allgemein

Ein *Kistenbaum* ist ein *binärer Baum* von Kisten, indem alle Knoten gleichzeitig in ihren gemeinsamen Vorgänger passen.

Als einen *Kistensatz* bezeichne ich eine Menge von Wurzeln mehrerer Kistenbäumen.

Mit $\text{elems}(ks)$ sei die Vereinigung der Menge aller Kisten aus den Kistenbäumen bezeichnet.

Das Grundproblem ist nun, zu einer Menge gegebener Kisten k_1, k_2, \dots, k_n ein Kistensatz ks mit den Wurzeln w_1, \dots, w_m zu erzeugen mit $\text{elems}(ks) = \{k_1, k_2, \dots, k_n\}$. Seien v_1, \dots, v_m die jeweiligen Volumina der Wurzeln w_1, \dots, w_m . Dann gilt es als weitere Aufgabe $\sum_{i=1}^m v_i$ zu minimieren.

1.2. Bruteforce

Die Liste wird entsprechend dem Volumen von groß nach klein sortiert. Die Liste wird nacheinander zu Kartonsätzen kombiniert. Eine Hilfsfunktion erzeugt aus einer Menge von Kartonsätzen durch hinzufügen einer gegebenen Kiste die Menge aller möglichen Kistensätze. Diese werden dann weiter mit dem nächsten zu noch mehr Kistensätzen kombiniert. Am Ende sind alle Elemente der Liste abgearbeitet.

Im Folgenden ist der Algorithmus in Scala Code dargestellt. Ich habe mich bewusst gegen Pseudo-Code Notation entschieden. Der Scala Code ist meiner Meinung nach ebenso effektiv wie Pseudo-Code. Lediglich wenige Elemente funktionaler und objektorientierter Elemente müssen dem Leser bekannt sein.⁶

Wichtig ist, um den Code zu verstehen, dass $(\text{satz} ++< \text{kiste})$ alle Möglichkeiten erzeugt, wie man die Kiste in einen Satz einfügen kann.⁷

```

1  // Nacheinander die Kisten "auffalten" mit Hilfe der hilfsPacken Funktion
2  def packe = (Set[KistenSatz]() /: kisten) ( hilfsPacken )
3  def hilfsPacken(sätze: Set[KistenSatz], kiste: Kiste) =
4      if(sätze.isEmpty)
5          Set(KistenSatz(kiste :: Nil)) // KistenSatz nur mit der Kiste
6      else
7          (Set[KistenSatz]() /: sätze) { // Beginne mit leerer Menge
8              (menge, satz) =>
9                  menge ++ // Füge neue Möglichkeiten der menge hinzu
10                 (satz ++< kiste) // Erzeugt neue Möglichkeiten
11          }

```

Laufzeitverhalten Das Laufzeitverhalten dieses Algorithmus ist fatal. Es muss im letzten Schritt eine Kiste in bis zu $(n-1)!$ Kistensätze gepackt werden. Die Laufzeit eine Kiste in einen Kistensatz zu packen ist $O(n)$, es muss für jede Kiste des Kistensatzes Möglichkeiten erzeugt werden. Wir erhalten also $n! + n \cdot (n-1)! + (n-1) \cdot (n-2)! + \dots + 2 \cdot 1! + 1$. Sprich $O(n \cdot n!)$.

Auch wenn der worst-case meist nicht erreicht wird, beispielsweise wenn es für 40 Kisten eher 20! Möglichkeiten gibt, würde die Berechnung aller Möglichkeiten bereits $8 \cdot 10^{15}$ Jahre brauchen.⁸

Unter der gleichen Annahme, zeigt sich, dass etwa 15! Operationen in einer Stunde ausgeführt werden können. Sprich es können $2 \cdot (15-1) = 28$ Kisten in allen Möglichkeiten gepackt werden. (Unter der Annahme es gibt etwa $x!$ Möglichkeiten für $2x$ Kisten⁹)

Verkürzung der Laufzeit Eine Überlegung war, das Laufzeitverhalten durch Parallelisierung zu verkürzen. Allerdings verspricht dies aufgrund der hohen Laufzeitkomplexität von $O(n \cdot n!)$ kaum Abhilfe. Selbst bei 100 Kernen, sprich einer hundertfachen Beschleunigung¹⁰ können gerade mal 17! Operationen ausgeführt werden. Das entspricht $2 \cdot (17-1) = 32$ Kisten. Es können also $14\%(32/28 = 1.1428\dots)$ mehr Kisten gepackt werden, was nicht nennenswert viel ist.

Es kommt also für Frau Y. somit nicht in Frage die Packberechnung beispielsweise auf eine Rechnerfarm zu migrieren.

⁶Beispielsweise sollten Sie wissen, wie foldLeft, currying, etc. funktioniert.

⁷Nähere Erläuterungen dazu später, für den Algorithmus ist dies nicht direkt relevant.

⁸Unter der Annahme dass das Prüfen und Hinzufügen einer Kiste in eine andere Kiste 1 ns dauert.

⁹Die Annahme zeigt sich als gar nicht so schlecht, wie man in ?? sieht

¹⁰Dies wird in der Praxis nie erreicht. Es muss immer ein gewisser Overhead für Synchronisation und Sequentielle Programmabläufe "geopfert" werden.

Problem Frau Y. hat also ihre mittlerweile 25 Kisten optimal packen können. Dadurch ist nun Platz in ihrem Keller frei geworden. Sie sieht es als ironisch an, dass sie genau deswegen nicht mehr Kisten in ihren Keller stellen kann, weil sie versucht den Platzverbrauch ihrer Kisten zu minimieren. Sprich, sie könnte beispielsweise eine Kiste direkt daneben stellen, die nicht mehr in die Berechnung einbezogen werden kann. Es ist offensichtlich, dass die ursprüngliche Motivation dadurch nicht erreicht wird. Wenn beispielsweise 200 Kisten gepackt werden sollen, bleiben 170 Kisten neben 30 optimal gepackten ungepackt.

Hierzu habe ich zwei Lösungsideen erstellt. Die grundlegende Motivation ist, die optimale Packung aufzugeben und stattdessen in menschlicher Zeit ¹¹ trotzdem eine gute Packung auch zu einer großen Menge von Kisten zu finden.

1.3. Bruteforce nach Aufteilung

Ansatz Eine Möglichkeit wäre, den Bruteforce Algorithmus immer auf 30 Kisten anzuwenden und danach diese Kistensätze nebeneinander zu stellen.

Wichtig ist hierbei, dass die Kisten sinnvoll aufgeteilt werden, so dass jeder Kistenhaufen kleinere und größere Kisten hat um eine hohe Packdichte zu erreichen. ...

Laufzeitverhalten Dieser Algorithmus ist streng polynomiell. Genauer gesagt kann er sogar in $O(n \log n)$ Zeit ausgeführt werden. Dies ergibt sich aus einer Laufzeitanalyse des Algorithmus.

Zunächst müssen die Kisten sortiert werden in $O(n \log n)$. Das Aufteilen der Kisten in Gruppen braucht $O(n)$. Man erhält also $\frac{n}{30}$ Gruppen. Eine dieser zu packen geschieht in konstanter Zeit. (Auch wenn der Bruteforce ursprünglich eine Komplexität von $O(m \cdot m!)$ besitzt, ist mit $m = 30$ seine Laufzeit $O(30 \cdot 30!) = O(1)$, also konstant.) Daraus ergibt sich letztendlich

$$O(n \log n + n + \frac{n}{30} \cdot 1) = O(n \log n)$$

.

1.4. Online Packer

Ansatz Ein etwas anderer Ansatz ist, Kistensätze inkrementiell zu erzeugen. Daher, es wird ein Algorithmus erfordert, welcher zu einem - mehr oder weniger gut - gepacktem Kistensatz und einer zu packenden Kiste ein neuen Kistensatz liefert, der, möglichst dicht gepackt, diese enthält. Eine weitere Beschränkung, die ich an den Algorithmus stelle, ist, dass er in höchstens $O(n)$ Zeit diesen neuen Kistensatz liefert. Hat man nun solch einen Algorithmus, lassen sich alle Kisten zusammen in $O(n \cdot n) = O(n^2)$ Zeit packen bei inkrementieller Kistensatzerzeugung.

Onlinealgorithmus Dieser Algorithmus kann von Frau Y. jedoch auch verwendet werden, um eine Kiste in der eine Lieferung verpackt war, in ihren vorhandenen Kistensatz hinzuzufügen. Es handelt sich also um einen Onlinealgorithmus. Die Entwicklung eines Onlinealgorithmus ist in der Aufgabenstellung weder explizit noch implizit gefordert. Es handelt sich also um eine eigenständige Erweiterung. Ich finde sie insofern sinnvoll, da sie ein Anwendungsfall direkt erfüllt, nämlich genau den, wenn Frau Y. eine einzelne Kiste erhält. Für Frau Y. ist es nun zwar möglich, einen neuen Kistensatz in linearer Zeit zu erhalten, aber es muss noch sichergestellt werden, dass auch ein möglicherweise nötiges Umpacken in linearer Zeit ausgeführt werden kann. Sprich, wir betrachten auch die "Laufzeit" Frau Y.'s und nicht die eines Computers. Für nachfolgende Strategien betrachte ich deswegen auch immer die Laufzeit für das Umpacken, welches nötig ist um die Kiste hinzuzufügen.

Strategien Es gibt unterschiedliche *Strategien* um einen Platz für eine Kiste in einem Kistensatz zu finden. Recht naheliegend sind unter anderem folgende.

¹¹menschliche Zeit = Lebenserwartung eines Menschen (75 Jahre)

FindeHalbleeren Findet eine KisteHalb die noch Platz für die neue Kiste bietet.

FindeGößerenLeeren Findet eine KisteLeer die noch Platz für die neue Kiste bietet.

FindeZwischenraum Findet eine Kiste die durch Umpacken einer Kind-Kiste in die neue Kiste genug Platz für die neue Kiste bietet.

FindeKleinereWurzel Findet eine Wurzel-Kiste die in die neue Kiste passt.

Offlinealgorithmus Werden die Kisten vor dem inkrementiellem Packen nach Volumen von groß nach klein sortiert, können Strategien 3 und 4 ohne Beschränkung weggelassen werden. Diese suchen nämlich Kisten, die kleiner sind als die hinzuzufügende, welche jedoch wegen der Sortierung nicht existieren können. Da jedoch zur Sortierung die Kisten bekannt sein müssen, handelt es sich nicht mehr um einen Online- sondern einen Offlinealgorithmus. Die Existenzberechtigung dieses Algorithmus ergibt sich aus der Tatsache, dass bessere Ergebnisse bei vorheriger Sortierung erhalten werden können als mit dem ursprünglichem Onlinealgorithmus. ¹²

2. Implementierung

Zunächst wurde ein Kern implementiert, welcher Kisten und Kistensätze sinnvoll abbildet und hilfreiche Funktionen zur Operation auf diesen bietet. Die Datentypen wurden als unveränderbare Objekte implementiert um die Algorithmen zu vereinfachen.

Kisten Es gibt drei Arten von Kisten: KisteLeer, KisteHalb und KisteVoll. Eine KisteLeer enthält keine weitere Kiste, eine KisteHalb enthält eine Kiste und eine KisteVoll enthält zwei. Es wurde zunächst ein **trait** Kiste implementiert, welches eine Anwendungsschnittstelle “nach außen” bietet und eine Schnittstelle “nach innen”, welche von den drei Unterklassen implementiert werden muss. Das **trait** wurde als **sealed** implementiert, das heißt, nur Typen in der gleichen Datei dürfen dieses **trait** implementieren. Dies ermöglicht bessere Compiler-unterstützung bei Pattern-matching, da bekannt ist, dass es nur genau 3 Unterklassen von Kiste gibt.

Neben der Implementierung von **val** hashCode: Int und **def** equals(Kiste): Kiste zur Verwendung als Hashkeys wurden auch oft verwendete anwendungsspezifische Methoden implementiert. Zum einen existiert die Methode **def** +<(Kiste): Set[Kiste] welche alle Möglichkeiten **eine andere** Kiste in den durch **diese** Kiste definierten Kistenbaum gepackt werden kann zurückliefert. Weitere Methoden erwähne ich bei Benutzung.

Kistensatz Da eine Kiste die Wurzel eines Kistenbaums ist und diesen repräsentiert, (Betrachten Sie eine KisteLeer als einen ein-elementigen Kistenbaum), muss ein Kistensatz lediglich Referenzen auf die einzelnen Kiste-Objekte speichern. Es ergeben sich für die Datenstruktur, diesen Kistenwald (Menge von Kistenbäumen) zu verwalten, folgende drei Voraussetzungen.

1. Das Ersetzen eines [Teil]baumes sollte möglichst billig sein. Da das Ersetzen einer Kiste als Löschen und anschließendes Hinzufügen dieser in den Baum implementiert ist, müssen sowohl die Lösch- als auch Hinzufügefunktionen schnelle Laufzeiten haben. Da die Datenstruktur unveränderbar ist, liefert jede Veränderung in dem durch eine Kiste repräsentiertem Kistenbaum eine neues Objekt zurück. Dieses Objekt muss dann durch eine Lösch- und eine Hinzufügeoperation in den Baum des KistenSatz aktualisiert werden.
2. “Duplikate” müssen zugelassen sein, da es passieren kann, dass zwei Kisten genau die gleichen Maße haben.
3. Die Kistenbäume eines Kistensatzes müssen so sortiert sein, dass ein Vergleichen nach Elementen billig ist.

Diese drei Voraussetzungen erfüllt meiner Ansicht nach ein geordneter Binärbaum am besten. Es wurde also die Scala Standardklasse TreeMap[Kiste,Int] verwendet. Sie bildet jeweils eine Kiste k auf eine Zahl $i_k > 1$ ab. Diese Zahlen sind gleich der Anzahl der einzelnen Kiste (bzw. Kistenbaum) in diesem Kistensatz. Somit erfolgt eine hinzufügen einer Kiste k (Scala: $+(Kiste):Kistensatz$) mit dem Hinzufügen von $k \rightarrow 1$ in den Binärbaum, bzw. wenn k schon erhalten ist, mit dem Setzen von $k \rightarrow i_k + 1$. Analog erfolgt auch das Entfernen einer Kiste k (Scala: $-(Kiste):Kistensatz$), daher wenn k nicht enthalten oder $i_k = 1$ entferne k aus dem Binärbaum, andernfalls setze $k \rightarrow i_k - 1$.

¹²Siehe auch: 3.1

Kistenpacker Da eine Menge von verschiedenen Algorithmen entwickelt wurden, bietet es sich an, von mehreren Algorithmen verwendete Funktionen auszulagern. Dies erfolgte in Scala durch Verwendung einer **trait**-Hierarchie.

3. Programmabläufe

3.1. Algorithm-Contest (Packdichte)

3.2. Algorithm-Contest (Laufzeit)

4. Programmtext

5. Programmnutzung

C. Zweite bearbeitete Aufgabe: (2) Containerklamüsel

1. Lösungsidee

1.1. Vorüberlegungen

1.2. Algorithmen

Ergebnisoptimaler Algorithmus

Optimale Ergebnisse

Laufzeitverhalten

Optimaler Algorithmus

Optimale Ergebnisse

Optimale Laufzeitkomplexität

2. Implementation