Anmerkungen zu Übungsblatt 11
R-Bäume

Jürgen Kalinski
Universität Bonn, Institut für Informatik III
Römerstr. 164, 53117 Bonn
cully@cs.uni-bonn.de

10. Februar 1999

Übungsblatt 11, Aufgabe 1

Die Struktur von R-Bäumen orientiert sich an den in der Vorlesung und den Übungen ausgiebig diskutierten B-Bäumen und B^+-Bäumen.

1. Knoten des Baumes sind mit Seiten des Hintergrundspeichers assoziiert.
2. Der Baum ist ausgeglichen.

Während B- und B^+-Bäume Suchbäume für eindimensionale Daten sind, verwalten R-Bäume zweidimensionale Daten, genauer gesagt: achsenparallele Rechtecke. Jeder Knoten im R-Baum speichert die Koordinaten von maximal M Rechtecken. Die Vater-Sohn-Eigenschaft des Baumes beruht auf folgender neuer Idee: Wenn ein Knoten die Koordinaten von n Rechtecken $r_1, \ldots r_n$ speichert, wird der Knoten in seinem Vaterknoten durch das kleinste r_1 bis r_n umfassende Rechteck repräsentiert. Dieser Repräsentant ist im Vaterknoten mit einem Zeiger auf den Sohn assoziiert. Mit anderen Worten: Der Aufstieg im Baum geht mit zunehmend gröberen Umschreibungen einher, der Abstieg im Baum mit Verfeinerungen (siehe Abb. 1).

Betrachten wir nun die Anfrage nach allen im R-Baum gespeicherten Objekten, die das Rechteck r_a mit den Koordinaten $(40, 60)$ und $(60, 90)$ schneiden. Dem Wurzelknoten läßt sich entnehmen, daß r_a sich mit 14 und 235 schneidet. Dem entnehmen wird, daß die mit diesen Rechtecken assoziierten Teilbäume möglicherweise r_a schneidene Rechtecke enthalten. Dahingegen kann der zweite Blattknoten ignoriert werden. Nicht einmal seine Vergröberung 678 schneidet r_a. Wir sind hier also gezwungen, sowohl in den linken wie in den rechten Teilbaum hinabzusteigen, um dort die Suche rekursiv fortzusetzen. Der Abstieg in den linken Teilbaum erweist sich allerdings als vergebliche Mühe. Die detaillierte Objektgeometrie ergibt, daß hier kein r_a schneidendes Rechteck existiert. (Beachte: Trotz der Ausgeglichenheit des R-Baums kann hier somit im Gegensatz zum B-Baum keine logarithmische Zugriffszeit garantiert werden.)

Betrachten wir nun, wie ein R-Baum aufgebaut werden kann. Wir nehmen an, daß die ersten drei in der Übungsaufgabe benannten Rechtecke nacheinander in einen anfangs leeren Baum eingefügt werden. Mit dem vierten Rechteck kommt es zu einem Overflow, da der Knoten nur maximal drei Rechtecke aufnehmen kann. Wie beim B-Baum kommt es nun zum Splitting. Ein neuer Knoten wird erzeugt, die Rechtecke werden auf beide Knoten verteilt und umschreibende Repräsentanten für beide Knoten dem Vaterknoten hinzugefügt (siehe Abb. 2).

Die wesentliche Frage hierbei ist jedoch: In welcher Weise sollen die Rechtecke auf zwei Knoten verteilt werden? Vom Standpunkt anschließender Suchoperationen aus gesehen, sollten die umschreibenden Repräsentanten im Vaterknoten möglichst klein sein. Große (sprich: grobe) Repräsentanten führen mit größerer Wahrscheinlichkeit zu unnötigen Abstiegen in einen Teilbaum als kleine. Gesucht ist also diejenige Partition von n Rechtecken $r_1, \ldots r_n$ in zwei Mengen R_1 und

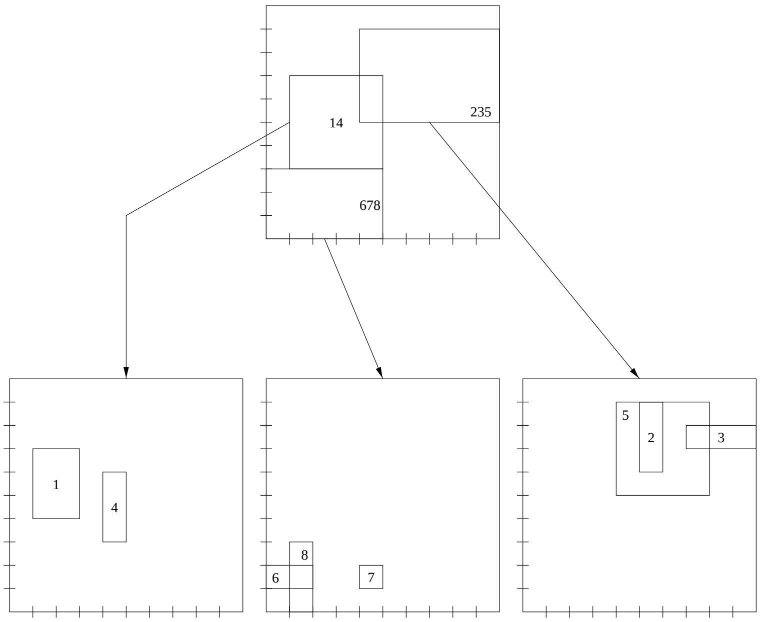

Abbildung 1: Ein R-Baum für $M = 3$

R_2, so daß die Flächensumme des R_1 umschreibenden und des R_2 umschreibenden Rechtecks minimal wird. Es sollte klar sein, daß für hinreichend große n dieses Problem nicht optimal gelöst werden kann: Es gibt 2^n Partitionen. Selbst wenn beide Mengen $n/2$ Objekte enthalten sollen, wären $\binom{n}{n/2}$ Partitionen (und damit immer noch exponentiell viele) zu inspizieren (siehe Abb. 3).

An dieser Stelle muß man also zwangsläufig mit Heuristiken an das Problem herangehen. Wir skizzieren im folgenden solche Heuristiken, auf deren Grundlage das Problem in zwei Schritten gelöst wird:

1. Bestimme zwei „Saat"-Rechtecke r_{i_1} und r_{i_2} als initiale Elemente für die Partitionen, d.h. $R_1 := \{r_{i_1}\}$ und $R_1 := \{r_{i_2}\}$.
2. Solange noch nicht alle Rechtecke einer der Mengen zugewiesen wurde, bestimme das nächste einzuordnende Rechteck und weise es einer der Mengen zu.

Die hier verwendeten Heuristiken sind nur noch von quadratischer Komplexität.

Schritt 1: Die Idee hinter der initialen Zuweisung lautet:

> „Bestimme diejenigen Rechtecke r_{i_1} und r_{i_2}, die man am wenigsten innerhalb derselben Menge sehen will."

Mit der ursprünglichen Zielsetzung der Flächenminimierung vor Augen, kann man hier also diejenigen zwei Rechtecke auswählen, die zusammengenommen den größten Verschnitt aufweisen. Genauer gesagt: Sei R eine Menge von rechtecken und $F(R)$ die Fläche des kleinsten alle $r \in R$ umschreibenden Rechtecks, so suchen wir dasjenige Paar, bei dem der Verschnitt

$$F(\{r_{i_1}, r_{i_2}\}) - F(\{r_{i_1}\}) - F(\{r_{i_2}\})$$

maximal ist. Hierzu müssen also alle $\binom{n}{2}$ Rechteckspaare betrachtet werden.

Schritt 2: Angenommen, es sind noch m Rechtecke übrig, die noch keiner der beiden Mengen zugewiesen wurden. Die Idee lautet nun:

> „Bestimme dasjenige Rechteck mit der deutlichsten Neigung zu einer der Mengen R_1 bzw. R_2."

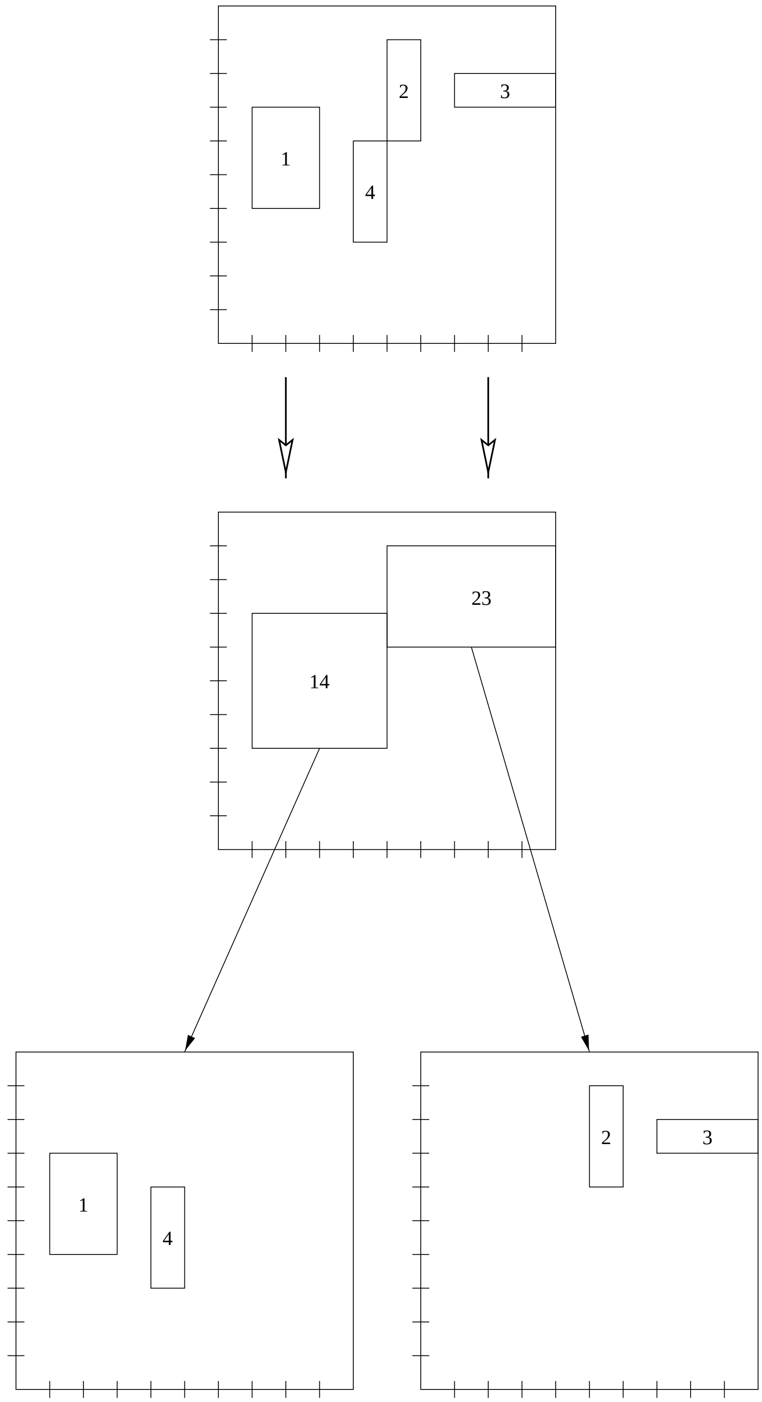

Abbildung 2: Splitting im R-Baum

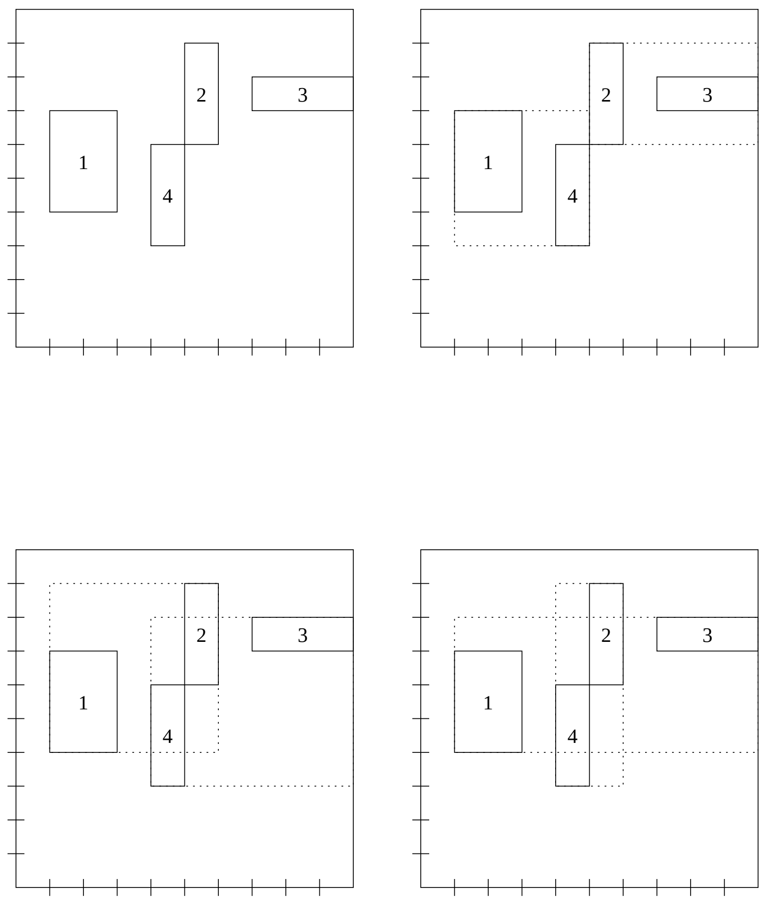

Abbildung 3: Einige Partitionierungsmöglichkeiten

Wiederum die Flächenminimierung im Hintergrund wählen wir dasjenige Rechteck r_{i_j}, bei dem die Flächenzuwächse, die r_{i_j} in R_1 bzw. R_2 herbeiführen würde, am weitesten auseinanderliegen. Genauer: Bestimme dasjenige r_{i_j} bei dem

$$|\, F(R_1 \cup \{r_{i_j}\}) \; - \; F(R_2 \cup \{r_{i_j}\}) \,|$$

maximal ist und weise es derjenigen Menge zu, bei der es zu dem geringeren Flächenzuwachs führt.

Nähere Ausführungen findet der interessierte Leser in Antonin Guttmans Paper (siehe [1]).

Literatur

[1] GUTTMAN, ANTONIN: *A Dynamic Index Structure for Spatial Searching.* Proc. of the ACM SIGMOD Conference. SIGMOD Record, 14(2):47–57, 1984.