

Aufgabenblatt 2: Indexstrukturen

- Abgabetermin: **Dienstag, 24.11.09 (23:59 Uhr)**
- Zur Prüfungszulassung muss ein Aufgabenblatt mit mind. 25% der Punkte bewertet werden und alle weiteren Aufgabenblätter mit mindestens 50% der Punkte.
- Die Aufgaben sollen in Zweiergruppen bearbeitet werden.
- Abgabe:
 - per E-Mail an `dbs2-200910@hpi.uni-potsdam.de` mit Subject
Abgabe DBS II: Aufgabenblatt <n> Namen
 - ausschließlich pdf-Dateien
 - eine Datei pro Aufgabe mit folgendem Dateinamen:
`blatt<aufgabenblattNr>aufgabe<aufgabenNr><Nachnamen>.pdf`
Bitte **keine Leerzeichen, Unterstriche, Umlaute, Sonderzeichen, ...** im Dateinamen!
 - **jedes Blatt beschriftet mit Namen**
 - Wir korrigieren die Abgaben aufgabenweise. Das beschriebene Verfahren vereinfacht uns die Arbeit erheblich!

Aufgabe 1: Dichtbesetzte und dünnbesetzte Indexe

Betrachte Blöcke, die entweder 30 Datensätze oder 200 Schlüssel-Pointer-Paare aufnehmen können. Die Blöcke dürfen nur zu 80 % gefüllt werden. Wie viele Blöcke werden in Abhängigkeit der Anzahl der Datensätze n benötigt für

- a) die Datensätze und einen dichtbesetzten Index? **2 P**
- b) die Datensätze und einen dünnbesetzten Index? **2 P**

Aufgabe 2: Sekundärindexe

Gegeben sei die Relation `Filme(Titel, Jahr, Länge, inFarbe, Studioname, Produzent)` (analog zur VL). Die Relation enthält 105 Filme des Studios Babelsberg. Weiterhin sind 215 Filme aus dem Jahr 2007 gespeichert, wobei nur einer von diesen in Babelsberg aufgenommen wurde.

Berechne jeweils die maximal notwendige Anzahl der I/O-Operationen bei der Beantwortung der folgenden Anfrage unter den im Folgenden genannten Rahmenbedingungen:

```
SELECT *  
FROM Filme  
WHERE Studioname = 'Babelsberg' AND Jahr = 2007
```

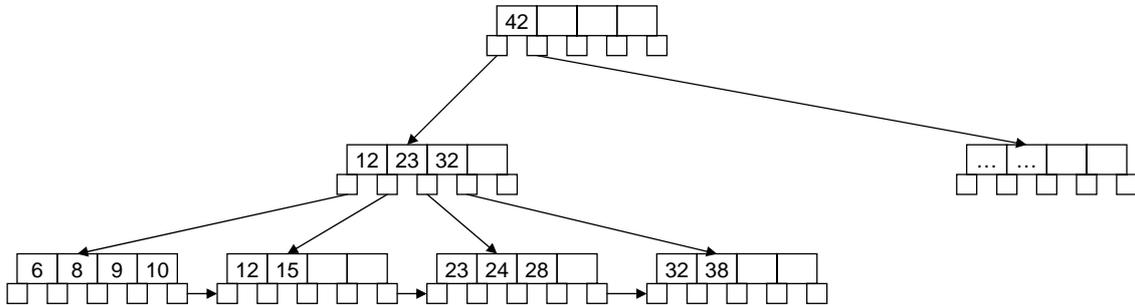
Ein Block soll entweder drei Records oder 10 Schlüssel-Zeiger-Paare oder 50 Zeiger enthalten. Nimm für die Berechnung an, dass die gegebenen Sekundärindexe bereits vollständig im Hauptspeicher geladen sind (nicht aber die Buckets und Records).

Gegeben seien die folgenden Sekundärindexe. Berechne jeweils die *maximal* notwendige Anzahl der I/O-Operationen. Stelle deine Überlegungen nachvollziehbar dar.

- a) Es existiert jeweils ein Sekundärindex auf `Studioname` und `Jahr`. Für beide Indexe werden Buckets (Indirektion) verwendet. **3 P**
- b) Es existiert ein Sekundärindex auf `Studioname` für den keine Buckets (Indirektion) benutzt werden. Auf `Jahr` existiert kein Index. **3 P**
- c) Es existiert ein Sekundärindex auf `Jahr` für den keine Buckets (Indirektion) benutzt werden. Auf `Studioname` existiert kein Index. **3 P**

Aufgabe 3: Einfügen und Löschen in B^+ -Bäumen

Gegeben sei der folgende B^+ -Baum:



- Füge die Schlüssel 17, 18, 19 und 11 (in dieser Reihenfolge) ein. Gib den B^+ -Baum nach dem Einfügen der 19 und der 11 an. Beschreibe kurz für jeden Schlüssel die notwendigen Schritte zum Einfügen. **10 P**
- Lösche die Schlüssel 17, 18 und 23 (in dieser Reihenfolge). Gib den B^+ -Baum jeweils nach dem Löschen der 18 und 23 an. Beschreibe kurz für jeden Schlüssel die notwendigen Schritte zum Löschen. **10 P**

Aufgabe 4: korrekte B^+ -Bäume

Ist der folgende gegebene B^+ -Baum ein gültiger B^+ -Baum mit max. 4 Schlüssel(n) (und damit max. 5 Pointern) pro Knoten? Falls nein, nenne alle Verletzungen der B^+ -Baum-Definition und begründe sie kurz. **10 P**

