



8. Übungsblatt

(Testwoche: 8. - 10. Juni 2010)

Einführung in Datenbanksysteme
Datenbanken für die Bioinformatik

Heinz Schweppe, Katharina Hahn

Aufgabe 1 (B-Bäume)

4 Punkte

Betrachten Sie ein DBMS mit folgenden technischen Daten:

- Die Blockgröße ist 2KB (also 2048 Byte).
- Verweise auf Datensätze bzw. Knoten sind 12 Byte groß.
- Ein Blockheader ist 56 Byte groß.

Es existiert ein B-Baum, bei der der Suchschlüssel 8 Byte groß ist. Bestimmen Sie die Höhe des B-Baums, wenn insgesamt 1 000 000 Datensätze in diesem indexiert sind. Geben Sie ihren Rechenweg an.

Aufgabe 2 (Indexierung, B-Bäume)

4 Punkte

In einer Datenbank wird eine typische Anfrage auf eine Tabelle häufig durchgeführt. Diese Anfrage liefert durchschnittlich 100 Elemente zurück. Die verwendete Festplatte hat eine maximale Zugriffszeit (seek + latency) von 8ms und einen Datendurchsatz von mindestens 15MB/s. Die Blockgröße beträgt 8KB und jeder Block ist zu mindestens 80% gefüllt.

Berechnen Sie, ab welcher Gesamtanzahl von Datensätzen in der Tabelle die Benutzung eines B-Baumes effizienter ist als eine sequentielle Suche. Schätzen Sie dabei alle Berechnungen nach oben ab und gehen Sie von einer Höhe des Baumes von 3 aus.

Aufgabe 3 (Konfliktserialisierbarkeit)

(3+3+3) Punkte

Gegeben sind die drei Schedules für die Transaktionen T1 und T2 (sowie T3 in Aufgabenteil c). Überprüfen Sie, ob diese serialisierbar sind. Geben Sie alle **Konfliktpaare** an und zeichnen Sie jeweils den entsprechenden **Konfliktgraphen** auf. Geben Sie für jeden konfliktserialisierbaren Ausführungsplan (Schedule) einen **äquivalenten seriellen Plan** an, z.B. T3 → T1 → T2.

- r1(a), w1(a), r1(b), w1(b), r2(b), w2(b), r2(c), w1(c)
- r1(c), r2(b), w1(a), w2(b), r1(b), r2(c), w1(b), w2(c)
- r1(a), r2(b), r3(a), w3(a), r1(b), w1(b), r2(c), r1(c), r2(d), r1(d), w1(d), w3(c)

Aufgabe 4 (Konfliktserialisierbarkeit)

5 Punkte

Zeigen Sie, dass es serialisierbare Historien gibt, die ein Scheduler auf Basis des 2PL-Protokolls nicht zulassen würde. Zeigen Sie also, dass die Klasse der serialisierbaren Historien größer ist als die Klasse aller nach dem 2PL-Protokoll generierbarer Historien.