



Universität Freiburg  
Institut für Informatik  
Prof. Dr. G. Lausen  
Dr.-Ing. T. Hornung

Georges-Köhler Allee, Geb. 51  
D-79110 Freiburg  
lausen@informatik.uni-freiburg.de  
hornungt@informatik.uni-freiburg.de

Übungen zur Vorlesung  
*Datenbanken und Informationssysteme*  
Wintersemester 2012/2013  
30.1.2013

## 13. Aufgabenblatt: Physische Datenorganisation & Auswertung von Anfrageoperatoren

### Aufgaben, die nicht bewertet werden

#### Übung 1

Sei  $Sort()$  ein Verfahren zum Sortieren innerhalb des Internspeichers.

- (a) Skizzieren Sie unter Verwendung von  $Sort()$  ein externes Sortierverfahren  $extSort()$ , mittels dem eine beliebige extern abgelegte Relation sortiert werden kann. Verfolgen Sie dabei die Strategie, zunächst Ausschnitte der Relation zu sortieren (analog zu *Merge Sort*), die jeweils vollständig in den Primärspeicher geladen werden können.
- (b) Angenommen,  $extSort()$  stehen drei Seiten des Datenbankpuffers zur Verfügung, um Seiten der Relation zu lesen, bzw. Seiten zu schreiben. Die Seiten enthalten gerade jeweils 2 Tupel.  
Demonstrieren Sie das Verfahren für die folgende Eingabe bestehend aus den folgenden 7 Seiten:

[3, 4], [6, 2], [9, 4], [8, 7], [5, 6], [3, 1], [2]

- (c) Sei  $N$  die Anzahl Seiten der Relation. Es stehen  $k$  Seiten im Datenbankpuffer zum Sortieren zur Verfügung stehen. Geben Sie für  $N = 10^6$  die Anzahl benötigter Seitenzugriffe an. Was ergibt sich für  $k = 100$ ?

### Aufgaben, die bewertet werden

#### Übung 2 (2 Punkte)

Es soll der natürliche Verbund der Relationen zweier Relationenschemata  $R(A, B)$  und  $S(A, C)$  mittels des Sort-Merge Verfahrens berechnet werden. Sie können annehmen, dass beide Relationen bzgl. der Verbund-Attribute bereits aufsteigend sortiert sind.

Es gelte die Restriktion, dass zu jedem Zeitpunkt nur eine Seite von  $R$  und eine Seite von  $S$  im Hauptspeicher gehalten werden dann.

Sei  $n$  die Anzahl Tupel von  $R$  und  $N$  die Anzahl Seiten von  $R$ , bzw.  $m$  die Anzahl Tupel von  $S$  und  $M$  die Anzahl Seiten von  $S$ . Es soll die für die Verbundberechnung benötigte Anzahl Seitenzugriffe berechnet werden. Für das Schreiben des Endresultates erforderliche Seitenzugriffe sollen ignoriert werden.

- (a) Wie viele Seitenzugriffe werden im worst-case benötigt, wenn die Verbund-Attribute in beiden Relationen Schlüssel sind?
- (b) Wie viele Seitenzugriffe werden im worst-case benötigt, wenn die Verbund-Attribute in beiden Relationen kein Schlüssel sind?

### Übung 3 (4 Punkte)

Betrachten Sie den natürlichen Verbund zweier Relationen zu den Schemata  $R(A, B)$  und  $S(B, C)$ . Angenommen ein Tupel zu  $R$  benötige 50 Byte und ein Tupel zu  $S$  100 Byte. Die Seitengröße betrage 4 kB, die Relation zu  $R$  benötige 10000 Seiten und die Relation zu  $S$  5000 Seiten. Die Blockgröße betrage 100 Seiten und der Datenbankpuffer habe die Größe 102 Seiten. Diskutieren Sie für die unterschiedlichen Strategien

- Block-Nested-Loop-Verbund,
- Index-Nested-Loop-Verbund,
- Sort-Merge-Verbund und
- Hash-Verbund

den bei Auswertung des Verbundes entstehenden Aufwand.

Abzugeben durch Einwurf in den Briefkasten Raum 01-025 Gebäude 51 bis spätestens 07.02.2013, 12:00 Uhr