



Datenbanken II

Wintersemester 2017/18

Checkliste Quiz 3

Daniel Kocher, Willi Mann

January 24, 2018

Diese Checkliste soll den Studierenden bei der Quiz-Vorbereitung helfen. Es werden Themengebiete/Fragestellungen aufgezählt, die die Studierenden auf jeden Fall beherrschen sollen. Wir weisen allerdings darauf hin, dass die Bearbeitung der Checkliste alleine keine Garantie ist, dass das Quiz positiv bestanden wird. Sie stellt lediglich ein weiteres Hilfsmittel zur Vorbereitung dar.

Die genannten Themengebiete sind aus praktischer Sicht zu verstehen, d.h. beim Quiz werden praktische Beispiele gelöst (ähnlich wie im Proseminar).

1 Sortieren

Kapitel 3 aus dem dritten VO-Foliensatz *Anfragebearbeitung*¹.

- Unterschied zwischen logischer Sortierung (bspw. durch Sekundärindex) und physischer Sortierung (bspw. durch Clustered Index).
- Unterschied zwischen Hauptspeicher-/Platten-Sortieralgorithmen und wann welche Form angewendet werden kann.
- Externes Merge-Sort:
 - Grundidee
 - Unterschied *Create-Runs-Schritt* und *Merge-Schritt*.
 - Gegeben sei eine Puffergröße $M \geq 3$ und eine Anzahl an Tupel pro Block. Ausführen des Algorithmus (Läufe erzeugen und die einzelnen Merge-Schritt zeichnen bis zum Endergebnis). Siehe auch Aufgaben aus dem Proseminar.
 - Gegeben sei eine Anzahl an Tupel und eine Anzahl an Tupel pro Block. Finden der minimalen Puffergröße M , sodass nicht mehr als 1 Merge-Schritt benötigt wird. Siehe auch Aufgabe aus dem Proseminar.
 - Kostenanalyse: Kosten nach der Formel und tatsächliche Kosten (d.h. tatsächliche Blockzugriffe zählen).

2 Selektion

Kapitel 4 aus dem dritten VO-Foliensatz *Anfragebearbeitung*¹.

- Grundstrategien für die Auswertung: Unterschied sequentielles Lesen der Datei und Index-Suche.
- Verschiedene Arten von Prädikaten: Gleichheits- bzw. Bereichsanfrage, konjunktive/disjunktive Selektion.
- Verschiedene Auswertungsstrategien ($A1 - A7$).

¹<https://dbresearch.uni-salzburg.at/teaching/2017ws/db2/db2.03-handout-1x1-prev.pdf>

- Implikationen der verschiedenen Indexarten.
- Bitmap Index Scan und wann macht es Sinn diesen anzuwenden?
- Anhand eines Beispiels die effizienteste Auswertungsstrategie für eine gegebenen Selektion bestimmen (Relationen und Eigenschaften sowie vorhandene Indizes gegeben). Siehe auch Aufgaben aus dem Proseminar.

3 Join

Kapitel 5 aus dem dritten VO-Foliensatz *Anfragebearbeitung* ¹.

- Verschiedene Join-Arten (Theta-/Equi-Join, natürlicher Join).
- Was ist die Join-Selektivität, wie wird diese bestimmt und wann ist die Selektivität eines Joins schwach bzw. stark?
- Verschiedene Join-Algorithmen:
 - Nested-Loop-Join (NLJ)
 - Block-Nested-Loop-Join (BNLJ) mit und ohne Zick-Zack-Modus
 - Indexed-Nested-Loop-Join (INLJ)
 - Merge-Join (MJ) bzw. Sort-Merge-Join (SMJ)
 - Hash-Join (HJ)
- Voraussetzungen/Implikationen der einzelnen Join-Algorithmen.
- Welcher Join-Algorithmus kann für welche Join-Arten angewendet werden?
- Berechnung der Kosten der einzelnen Join-Algorithmen.
- Ausführen der einzelnen Join-Algorithmen anhand eines Beispiels. Vgl. Hash-Join-Aufgabe aus dem Proseminar.
- Innere bzw. äußere Relation für die einzelnen Join-Algorithmen bestimmen, sodass die Join-Kosten minimal sind.
- BNLJ mit Zick-Zack-Modus: Heuristik zur optimalen Ausnutzung des Puffers.
- INLJ: Kosten des Index-Lookups berechnen für statischen/dynamischen Hash-Index (bspw. mit Overflow Buckets) oder B⁺-Baum-Index. Siehe auch Aufgaben aus dem Proseminar.