

Staatsexamen 66116 / 2019 / Frühjahr

Thema 2 / Teilaufgabe 2 / Aufgabe 6

(Vermischte Datenbank-Fragen)

Stichwörter: Datenbank-Übersicht, Natural-Join, Equi-Join, Theta-Join, UNION, INTERSECT, EXCEPT

Begründen oder erläutern Sie Ihre Antworten.

- (a) Erklären Sie kurz den Unterschied zwischen einem Natural-Join und einem Equi-Join.

Ein Natural Join ist eine Kombination von zwei Tabellen, in denen Spalten gleichen Namens existieren. Die Werte in diesen Spalten werden sodann auf Übereinstimmungen geprüft (analog Equi-Join). Einige Datenbanksysteme erkennen das Schlüsselwort NATURAL und eliminieren entsprechend automatisch doppelte Spalten.

Während beim Kreuzprodukt keinerlei Anforderungen an die Kombination der Datensätze gestellt werden, führt der Equi-Join eine solche ein: Die Gleichheit von zwei Spalten.

^a^awiki.selfhtml.org

- (b) Erläutern Sie kurz was man unter einem Theta-Join versteht.

Ein Theta-Join ist eine Verbindung von Relationen bezüglich beliebiger Attribute und mit einem Selektionsprädikat. ^a

^a<https://www.datenbank-grundlagen.de/theta-join.html>

- (c) Was versteht man unter Unionkompatibilität? Nennen Sie drei SQL-Operatoren welche Unionkompatibilität voraussetzen.

Bestimmte Operationen der relationalen Algebra wie Vereinigung, Schnitt und Differenz verlangen Unionkompatibilität. Unionkompatibilität ist eine Eigenschaft des Schemas einer Relation. Zwei Relationen R und S sind genau dann union-kompatibel, wenn folgende Bedingungen erfüllt sind:

- (i) Die Relationen R und S besitzen dieselbe Stelligkeit n , d. h. sie haben die selbe Anzahl von Spalten.
- (ii) Für alle Spalten der Relationen gilt, dass die Domäne der i -ten Spalte der Relation R mit dem Typ der i -ten Spalte der Relation S übereinstimmt ($0 < i < n$).

Die Namen der Attribute spielen dabei keine Rolle. ^a

SQL-Operatoren mit Unionkompatibilität

- UNION
- INTERSECT
- EXCEPT

^a<https://studylibde.com/doc/1441274/übungstool-für-relationale-algebra>

- (d) Erläutern Sie Backward und Forward Recovery und grenzen Sie diese voneinander ab.
- (e) Erklären Sie das Zwei-Phasen-Freigabe-Protokoll.
- (f) Erläutern Sie Partial Undo / Redo und Global Undo / Redo und deren Bedeutung für die Umsetzung des ACID-Prinzips. Geben Sie zu jeder dieser Konzepte an, ob System-, Programm- oder Gerätefehler damit korrigiert werden können.
- (g) Erklären Sie das WAL-Prinzip (Write ahead logging)!

Das sogenannte write ahead logging (WAL) ist ein Verfahren der Datenbanktechnologie, das zur Gewährleistung der Atomarität und Dauerhaftigkeit von Transaktionen beiträgt. Es besagt, dass Modifikationen vor dem eigentlichen Schreiben (dem Einbringen in die Datenbank) protokolliert werden müssen.

Durch das WAL-Prinzip wird ein sogenanntes „update-in-place“ ermöglicht, d. h. die alte Version eines Datensatzes wird durch die neue Version an gleicher Stelle überschrieben. Das hat vor allem den Vorteil, dass Indexstrukturen bei Änderungsoperationen nicht mit aktualisiert werden müssen, weil die geänderten Datensätze immer noch an der gleichen Stelle zu finden sind. Die vorherige Protokollierung einer Änderung ist erforderlich, um im Fehlerfall die Wiederholbarkeit der Änderung sicherstellen zu können.

(h) Erklären Sie den Begriff „Datenbankindex“ und nennen Sie zwei häufige Arten.

Ein Datenbankindex ist eine von der Datenstruktur getrennte Indexstruktur in einer Datenbank, die die Suche und das Sortieren nach bestimmten Feldern beschleunigt.

Gruppierte Indizes (Clustered Index)

Bei der Verwendung eines gruppierten Index werden die Datensätze entsprechend der Sortierreihenfolge ihres Index-Schlüssels gespeichert. Wird für eine Tabelle beispielsweise eine Primärschlüssel-Spalte „NR“ angelegt, so stellt diese den Index-Schlüssel dar. Pro Tabelle kann nur ein gruppierter Index erstellt werden. Dieser kann jedoch aus mehreren Spalten zusammengesetzt sein.

Nicht-gruppierte Indizes (Nonclustered Index)

Besitzt eine Tabelle einen gruppierten Index, so können weitere nicht-gruppierte Indizes angelegt werden. Dabei zeigen die Einträge des Index auf den Speicherbereich des gesamten Datensatzes. Die Verwendung eines nicht-gruppierten Index bietet sich an, wenn regelmäßig nach bestimmten Werten in einer Spalte gesucht wird z.B. dem Namen eines Kunden. ^a

^a<https://www.datenbanken-verstehen.de/datenmodellierung/datenbank-index>