

Zentralübung ERDB

Programm heute:

- Übersicht Prüfungsstoff
- Fragerunde – Aber auch gerne während der Übersicht

Allgemeine Voraussetzungen

- Grundkenntnisse wichtig!
 - Bsp. Cube Operator
 - Bsp. Skyline in SQL
 - Bsp. Datalog in SQL ausdrücken
- Casts zwischen Datentypen nicht wichtig
- Folien mixen teilweise DB2 SQL Syntax und Postgre SQL Syntax → Beide sind in der Prüfung gültig, innerhalb einer Aufgabe muss aber konsistent sein

- Universitätsschema muss nicht auswendig gelernt werden
- Schema ist immer angegeben
- Teilw. auch Beispieldaten

Überblick Prüfungsstoff

Disclaimer: Werde heute hauptsächlich besprechen welche Themen besonders wichtig sind. Falls etwas auf den Folien nicht erwähnt ist, aber in den Übungen / Vorlesung besprochen wurde, kann darauß nicht geschlossen werden, dass es nicht in der Prüfung drankommt.

Recovery - **Alles wichtig, besonderer Fokus auf:**

- ACID
- Force / Steal
- Write Ahead Logging
- Log Struktur
- Physical / Logical logging
- Redo / Undo

Multiuser - **Alles wichtig mit Fokus auf**

- Welche Probleme können entstehen
- Serialisierbarkeit / Historien (auch die **Historienklassen und Zuordnung**)
- Deadlocks (Gründe / Erkennung / Vermeidung)
- 2PC / TS / Optimistic Concurrency Control
- ...
- Nicht wichtig:
 - Isolation Level in SQL
 - Synchronisation von Bäumen

Fokus auf Stoff in den Übungsblättern

- Angriffsarten
- k-Anonymität
- RSA (Verschlüsseln & Entschlüsseln, weshalb sicher?)
- SQL Injection
- SSL

- **Nicht so spannend**

- **Nicht so spannend**

- Datalog Theorie
 - Wann ist Programm sicher?
 - Stratifizierbar?
 - ...
- **Datalog Programme! Wichtig!**
 - Definition neuer Regeln
 - \neq , not(...), +
 - Einfache Regeln zu SQL übersetzen
 - Rekursion
 - **Keine Aggregationsfunktionen!**

- **Fragmentierung**
 - Vertikal / Horizontal
 - Korrektheit
 - Rekonstruktion
 - Join Auswertung
- Bloom Filter
- Chord
- Synchronisation: 2PC, Quorum,, ...

- *Cube Operator*
 - Bezug zu Group By
- **Top-K Anfragen**
 - Threshold / NRA
 - Skyline (in SQL mit und ohne „skyline“ operator)
- **Klassifikation**
 - Entscheidungsbaum
 - A Priori ←- **Nicht genau genug in den Unterlagen beschrieben**
 - **Assoziationsregeln** (Support / Konfidenz)

- **Speicherhierarchie**
- **Row vs Column Store**
- **Indexstrukturen**
 - ART

- Grundkenntnisse XML Schema
- **XPath**
 - Achsen
- **Xquery**
 - FLOWR
 - Falls Funktionen zur Lösung der Aufgaben benötigt werden, sind sie angegeben!

- Grundkenntnisse RDF / SPARQL
- Datenstromverarbeitung
- **PageRank**
 - PageRank per Hand berechnen
- **HITS**
- **TF-IDF**
 - Übung: TF-IDF per Hand berechnen

- **Schwierigkeit der Übungsblätter kann als Orientierung für die Prüfung genutzt werden**