Thema:	Prüfungsprotokoll Datenbanken
Datum:	13.01.2014
Prüfer:	Prof. Dr. Güting
Beisitzer:	Valdez
Note:	1,3

Kurs 1664 (Implementierungskonzepte für Datenbanken):

- Aufbau einer Datenbank (große Grafik) und alle Teile erläutern.
- Ablauf der Anfrageübersetzung (lexikalische, syntaktische und semantische Analyse) mit Erklärung
- Ablauf der Anfragoptimierung (3-teilige Grafik mit Erklärungen)
- Erklärung des Operatorbaums nur sehr grob
- Optimierung nur in groben Zügen, wie oben schon beschrieben

Kurs 1672 (Datenbanken II):

- Definition „Transaktion“
- 4 Probleme (lost update, inkonsistente Sicht, inkonsistente DB, Phantome)
- ACID und Erklärung der einzelnen Teile
- Definition „Serialisierbarkeit (gedachte serielle Ausführung + zyklenfreier Abhängigkeitsgraph)
- Erklärung des Abhängigkeitsgraphen
- Synchronisationsverfahren
- Arten von Sperren
- Sperrprotokolle (für pessimistische Verfahren)
- 2-Phasen-Sperrverfahren und die Verschärfungen (Preclaiming, strikte 2-Phasigkeit) erklären
- Auswirkungen auf Transaktionen erklären (fortgepflanzer Rollback, gleichzeitiger Datenbankzugriff)

Gedächtnisprotokoll 1664/1672

Datum: 10.10.2011

Prüfer: Herr Prof. Dr. Güting

Beisitzer: Herr Dr. Thomas Behr

Die Themen sollten meist kurz erläutert werden und dann gab es Detailfragen.

1664 Implementierungskonzepte für Datenbanksysteme

- Erklärung der Architektur eines DBS (stichpunktartig)
- Besonderheiten von geometrischen Daten erklären
- Indexstrukturen von geometrischen Daten erläutern
- Z-Ordnung Beispiel zeichnen und detailliert erklären
- R-Baum zeichnen und detailliert erklären

1672 Datenbanken II

- Probleme aufzählen
- Transaktionen definieren
- ACID-Prinzips
- Synchronisationsverfahren
- Sperrprotokoll und Verschärfungen

Gedächtnisprotokoll
21.11.2011, Prof. Dr. Güting.
Datenbanken II (1672) + Implementierungskonzepte für Datenbanksysteme (1664)
Dauer: ca. 30 Minuten.

1664:
DBMS "von unten nach oben": Alle Schichten und die einzelnen Bestandteile erklärt.
Der komplette Weg einer Anfrage: 3 Phasen: **Anfrageübersetzung**, -**Optimierung** und -**Ausführung**.
Jeweiligen Phasen ebenfalls erklärt:
Anfrageübersetzung: **lexikalische + syntaktische + semantische Analyse**
Anfrageoptimierung: **Anfrageumformung + Auswertungsplanerzeugung + Kostenschätzung**. Hier habe ich von mir aus erklärt, dass die Kostenschätzung eben nur eine "Schätzung" ist und auf Erfahrungswerten beruht, so dass der ermittelte günstigste Plan nicht unbedingt der tatsächlich günstigste Plan sein muss. Daraufhin kam die Frage, wieso das so sei. Ich habe für alle drei Teilbereiche der Kostenschätzung versucht, etwas zu erklären (Verbindungsprobleme bei der **Kommunikationszeit** vom/zum Anfragesteller, benötigte Daten könnten bereits im Hauptspeicher sein und müssten nicht erst noch aus Externspeicher geholt werden (**CPU**-**Rechenzeit** + **Zugriffszeit Externspeicher**). Das war zwar nicht ganz falsch, aber man wollte auf etwas anderes hinaus: Die Zwischenergebnisse von z.B. einem Join auf zwei Relationen mit jeweils zig-tausenden Datensätzen sind schwierig zu ermitteln.
Außerdem habe ich die drei Teilphasen der *Anfrageumformung* erklärt: **Anfrage-Umformung** (Beispiel a AND true = a etc.), -**Vereinfachung** (weshalb? Standardisierung! Wie? Eliminierung von Redundanzen!) -**Verbesserung** durch weitere Umformungen (wie? Operatorbaum erklärt. Selektionen nach unten. Wozu? Kleinere Mengen müssen bei binären Operationen verarbeitet werden.)
Operatorbaum - wie ist er aufgebaut? Innere Knoten = Operationen, Blätter = Relationen, innerer Knoten = Zwischenergebnis der Teilbäume, dessen Wurzel er ist.
Was ist ein Kartesisches Produkt gefolgt von einer Selektion im Operatorbaum? Ein Join.
Kann man eine Selektion weiter unten in den Baum direkt nach einer Relation stecken, also noch vor einem Join? Nur falls alle erwähnten Attribute der Selektion ausschließlich in einer der beiden per Join verbundenen Relationen vorhanden sind.

1672:
Was ist eine Transaktion? Folge von Lese/Schreiboperationen.
Wie erkennt man eine Transaktion? Begin of Transaction (BOT) + EOT-Marke. Alles dazwischen ist eine Transaktion.
Konzept der Transaktion? ACID erklären.
Welche Probleme können dabei auftreten? Lost Update, inkonsistente Sicht + DB, Phantome. Alles erklären.
Was bedeutet Serialisierbarkeit? **Serialisierbarkeitskriterium**? **Abhängigkeitsgraph**? Definitionen runter gerasselt.
Warum ist Lost Update kein Problem mehr bei serialisierter Schedule? Dazu habe ich einen Graphen zu W1(a),W2(a),R1(a),R2(a) (also gerade das Lost-Update Beispiel aus dem Skript) gezeichnet, der Zyklen enthielt. Genau das passiert eben nicht bei serialisierbarer Schedule.

Danach waren die knapp 30 Minuten auch schon zu Ende. Der DBS-Teil hat ca. 15 Minuten gedauert, DBII etwas weniger. Bei einigen Fragen habe ich nicht auf Anhieb verstanden, worauf Herr Prof. Güting hinaus wollte, er hat in diesen Situationen dann aber gleich Hilfestellung geleistet und mich auf den richtigen Pfad geleitet. Ich wusste leider nicht, dass im Operatorbaum eine Selektion nach einem kartesischen Produkt einen Join darstellt. Außerdem musste ich einmal bei der Kostenschätzung passen, s.o. Irgendwann sind wir auch kurz auf die konjunktive Normalform gekommen, ich kann dieses Thema allerdings keiner Frage mehr zuordnen. Hatte jedenfalls etwas mit Selektionen zu tun. Eine Folge von Selektionen auf eine Relation kann man auch als eine Konjunktion der Selektionen auf R beschreiben. Könnte auch zu den Anfrage-Umformungen passen.
Von meinem Vorgänger der Prüfung habe ich noch aufgeschnappt, dass die vier Joins abgefragt wurden. Nested-Loop Join (die vier foreach-Schleifen sollten sitzen), Index-Nested-Loop Join, Sort-Merge-Join, Simple-Hash-Join. Alle kurz erklären. Außerdem kam auch der Operatorbaum vor. Für mehr Smalltalk blieb keine Zeit mehr.

Insgesamt bin ich sehr zufrieden mit dem Prüfer. Fragen wurden auf ruhige Art gestellt und bei Verständnisfragen wurde ich direkt unterstützt. Als Prüfer kann ich Herrn Prof. Dr. Güting uneingeschränkt weiterempfehlen.
Viel Erfolg euch allen!

Gedächtnisprotokoll 1664/1672
Datum: 10.01.2011
Prüfer: Herr Prof. Dr. Güting
Beisitzer: Frau Simone Jandt

1664 Implementierungskonzepte für Datenbanksysteme

- Erklärung der Architektur eines DBS.
  Schichten und Komponenten jeweils mit kurzer Beschreibung.

- Welche Arten von Indexen gibt es?
  Klassifizierung: dünn / dicht, ..., jeweils mit Erklärungen.

- Besonderheiten von geometrischen Daten erklären.
  Komplex, von sehr großer und variabler Länge.

- Strategie der Indexstrukturen von geometrischen Daten erläutern.
  Zwei Schritte: Filterungsphase, Verfeinerungsphase.
  Indexstrukturen auf geometrische Objekte werden für Filterung verwendet und ermitteln eine Obermenge der Ergebnismenge.
  Geometrische Objekte (Linien, Regionen, nicht Punkte) werden durch ihre bounding boxes approximiert, welche als Suchschlüssel verwendet werden.

- Welche Indexstrukturen gibt es für Rechteck-Mengen?
  R-Bäume mit kurzer Erklärung.
  Ein R-Baum sollte als Beispiel gezeichnet werden (mit überlappenden Behälterregionen).

1672 Datenbanken II

- Probleme aufzählen und erklären.
  Lost update, inkonsistente Sicht, Inkonsistenz der Datenbank, Phantome.
  Habe jeweils ein Beispiel angegeben.

- Was sind Transaktionen?

- Erklärung des ACID-Prinzips.

- Synchronisationsverfahren erklären.
  Optimistische und pessimistische.

- Sperrprotokoll und Verschärfungen erklären.
  Zwei-Phasen-Sperrprotokoll (verhindert inkonsistente Sicht),
  strikte Zweiphasigkeit (verhindert fortgepflanzten Rollback),
  Preclaiming (verhindert Deadlock)

- Welche Sperrmodi gibt es?
  Lese-Sperren (ermöglicht konkurrierenden Lesezugriff), exklusive Sperren.

- Was kann man sperren?
  Tupel, Relationen, Seiten, Indexe, Datenbanken.

- Warum / wann macht es Sinn Tupel zu sperren?
  Andere Transaktionen können dann andere Tupel bearbeiten / lesen.

- Warum / wann macht es Sinn Relationen zu sperren?
  Bei Änderungen an viele oder allen Tupel der Relation, Schemaänderungen (Tabelle wird um Felder erweitert).

Prüfungsprotokoll für die Kurse 1664 (Implementierungskonzepte von Datenbanken) und 1672(Datenbanken II)
Datum: 14.4.2008
Prüfer: Prof. Güting

Kurs 1664

- Wie ist ein DBS aufgebaut? Aufzählung der einzelnen Schichten (Geräte- und Speichermanager, Systempuffermanager usw.)
- Recht allgemein gehaltene Erklärung der Aufgaben der einzelnen Schichten
- Welche Arten von Indizes unterscheidet man? Dicht-Dünn, Primär-Sekundär usw.
- Welche Indexstrukturen für mehrdimensionale Suchschlüssel gibt es? Für die Punktsuche ist Konkatenation der einzelnen Schlüssel möglich. Ablage der konkatenierten Schlüssel in B-Baum oder Hashen möglich. Wenn nicht nur Punktsuche unterstützt werden soll, sind Indexstrukturen für mehrdimensionale Punkte nötig. Hier habe ich zunächst die z-Ordnung genannt, Herr Güting wollte aber das Grid-File erklärt haben, das ich anhand einer Zeichnung erläutert habe.
- Erläuterung der z-Ordnung anhand einer Zeichnung. wofür stehen die einzelnen Bits der Codes, die den Zellen zugeordnet sind? Jeweils für die x- und y-Koordinate.
- Wie hilft die z-Ordnung, die Lage eines geometrischen Objekts zu bestimmen? Ablage der Codes der geschnittenen Zellen in einen B-Baum

Kurs 1672

- Welche Fehler können bei parallel arbeitenden Transaktionen entstehen? Lost update, inkonsistente Sicht, Inkonsistenz der DB, Phantome
- Was sind Phantome?
- Wie können die Fehler vermieden werden? Konzept der Transaktion und ACID Prinzip erklären
- Wann ist ein System von Transaktionen korrekt synchronisiert? Kriterium der Serialisierbarkeit und Erklärung des Abhängigkeitsgraphen.

Ich glaube, das war alles. Herr Güting ist als Prüfer sicherlich zu empfehlen, die Fragen sind zum Teil aber sehr spezifisch, so dass man sich eingehend mit dem Studienbriefen beschäftigen muß.

Viel Glück!

Prüfungsprotokoll
Typ: Videoprüfung
Datum: 02.05.2005
Prüfer: Prof. Güting
Kurs: 1664 (Implementierungskonzepte von Datenbanken) und 1672 (Datenbanken 2)

Fragen:

1664 (Implementierungskonzepte von Datenbanken)

- Welche Schritte werden durchlaufen, wenn eine Abfrage an die DB übergeben wird.
- Wie funktioniert die Abfrageoptimiertung
- Welche verschiedenen Implementierungen der Selektion gibt es
- Wann kann eine Selektion mit einem Kreuzprodukt vertauscht werden.

1672 (Datenbanken 2)

- Transaktionen: Was ist das
- ACID erklären
- Serialisierbarkeit
- Transaktionsverfahren (Optimistisch - Pessimistisch)
- Beschreiben von Lese- und Schreib-Lock
- Lockverfahren (2-Phasen, Preclaimbing, ...)
- Datenbankerweiterungen (Liste aufzählen)
- Multimedia Datenbanken genauer beschreiben

Wie alle anderen kann ich Prof. Güting uneingeschränkt weiterempfehlen. Und für alle die sich unter einer Videoprüfung nichts vorstellen können. Zum Einsatz kommt Netmeeting, eine Webcam und ein Headset.

Prüfungsprotokoll
Typ: mündliche Prüfung in Hagen
Datum: 20.04.2005
Prüfer: Prof. Güting
1. Teil einer Wahlfachprüfung
Kurs 1664 Implementierung von Datenbanksystemen
Note: unbenotet

Fragen:
Nennen Sie die einzelnen Komponenten eines Datenbanksystems „von unten nach oben“
>Damit waren die einzelnen Komponenten (Transaktionsmanager,
>Zugriffspfadmanager, Anfrageoptimierter etc) gemeint. Kurze Erklärung zu jeder
>Komponente. Z.B. Der Transaktionsmanager synchronisiert mehrere Transaktionen.

Was ist eine Anfrageoptimierung ?
>Ich habe kurz die einzelnen Punkte aufgezählt.
>Anfrageumformung, die sich wieder in Anfragestandardisierung, Anfragevereinfachung
>und Anfrageoptimierung unterteilt, danach die Kostenschätzung und die
>Auswertungsplanerzeugung.
>Darauf sind wir die einzelnen Punkte Schritt für Schritt durchgegangen. Herr Güting
>wollte also schon wissen, was in den einzelnen Phasen passiert. Ich habe zur
>Auswertungsplanerzeugung die Algorithmen erklärt (Iterative Improvement, Simulated
>Annealing und Two Phase Optimation)

Das war es schon. Zwar wurden nie Indexstrukturen oder „Externes Sortieren“ gefragt, allerdings baut Herr Güting diese Abfragen geschickt in die Themen ein. Er fragt dann z.B. zwischendurch mal, wie man denn die Selektion implementieren könnte. Generell sollte man sich mit allen im Script beschriebenen Themen auskennen.

Herr Güting ist ein ruhiger und netter Mensch, der es recht schnell schafft die Nervosität zu nehmen.

Viel Erfolg…