Datenbanken I

(Winter 2006/2007)

Prof. Dr. Stefan Brass

Institut für Informatik

Übungsleiterin: Dr. Annemarie Herrmann

Themen

- Grundlegende Datenbank-Begriffe, Motivation
- Etwas Logik
- Konzepte des relationalen Modells
- Die Datenbanksprache SQL (Schwerpunkt der Vorlesung)
- Einführung in DB-Entwurf (ER-Modell, BCNF, → DB IIA)
- Transaktionen (Recovery, Mehrbenutzerbetrieb)
- Sicherheits-Aspekte (Zugriffsrechte, Sichten)

  Weitere Themen sofern noch Zeit:
  Anwendungs-Programmierung (→ DB-Praktikum), QBE/Access,
  Objektrelationale DBen, Deduktive DBen, Suchmaschinen, XML.

Motivation (1)

Warum sind Datenbanken wichtig?

- „Informations-Gesellschaft“
- Information ist eines der wichtigsten Aktiva vieler Wirtschaftsunternehmen.
- Sie werden im Berufsleben in der einen oder anderen Form mit Datenbanken zu tun haben.
- Mit dem Boom des WWW gibt es noch mehr Anwendungen für Datenbanken.
- Datenbank-Experten sind (relativ) gesucht.

Motivation (2)

Warum machen Datenbanken Spaß?

- Interessante Verbindung von Theorie und Praxis, z.B. beruht SQL auf der Prädikatenlogik 1. Stufe.

  Viele DB-Begriffe sind so einfach, dass noch formal zu definieren.
- Zumeist sehr anschaulich, anwendungsorientiert.
- Relativ wenig komplexe Programmierung, wenn man nicht gerade ein neues DBMS implementiert.
- Interessante Algorithmen und Datenstrukturen, gute Anwendung für Parallelität.
- Noch nicht von Microsoft dominiert.

Motivation (3)

Aufgabe:

- Wo begegnen Ihnen im normalen Leben Datenbanken? Welche enthalten Informationen über Sie?
  - ◇ Supermarkt (Scannerkasse)
  - ◇ _____
  - ◇ _____
  - ◇ _____
  - ◇ _____
  - ◇ _____

Zeit und Ort (1)

Vorlesung (4 SWS):

- Mittwochs, 10^{15}–11^{45}, Raum 3.31.
- Freitags, 10^{15}–11^{45}, Raum 3.28.
- 12 LP im Grundstudium, 10 LP im Hauptstudium (für Wirtschaftsinformatiker) (inkl. Übung).

Anwesenheit nicht verpflichtend (Folien/Skript), aber:

- Eventuell wichtige Ankündigungen.
- Beispiele, kleine Aufgaben, Fragen der Hörer.
- Zeitaufwand für selbst nacharbeiten eher höher.

Zeit und Ort (2)

Übung (2 SWS):

- Tafelübung, Hausaufgaben besprechen (∼45 min), anschließend betreute Zeit am Rechner (∼45 min).
- Drei Gruppen (Anmeldung: Punkte-DB im Web):

ID	Tag	Zeit	Raum	Beginn
DI10	Dienstags	10^{15}–11^{45}	1.03/3.34	24.10.
MI14	Mittwoch	14^{15}–15^{45}	1.30/3.34	25.10.
DO16	Donnerstag	16^{15}–17^{45}	1.30/3.34	26.10.

Übungszeiten können sich noch ändern!

Übungsschein (1)

Klausur (voraussichtlich am Mittwoch, 14.02.2007):

- Bücher, Notizen, etc. können verwendet werden.
  Bücher sind nur nützlich wenn man sie vorher gelesen hat.
- Praktische Anwendung, kein Auswendiglernen.
  Ca. 50%: SQL-Anfragen.
- Alte Klausuren im Netz.

Gewichtung Hausaufgaben/Klausur (Maximum gilt):

Formel	Hausaufgaben	Klausur
I	30%	70%
II	0%	100%

Übungsschein (2)

Hausaufgaben (optional, dringend empfohlen):

- Hausaufgaben sind eine wichtige Vorbereitung auf die Klausur!
- Bearbeitung einzeln oder in Gruppen bis ca. 3–4 Personen.

  Es können nur Teilnehmer der gleichen Übungsgruppe (gleicher Termin) Hausaufgaben zusammen abgeben. Von gemeinsam erarbeiteten Lösungen bitte nur ein Exemplar abgeben (bzw. nur eine EMail)!
- Ausgabe Freitags in der Vorlesung bzw. im Netz, Abgabe Freitag der nächsten Woche in Vorlesung.

  Oder per Email (bis 12^{15}, bitte PDF): `herrmann@informatik.uni-halle.de`

Übungsschein (3)

Hausaufgaben, Fortsetzung:

- Jedes Gruppenmitglied muß jede von der Gruppe abgegebene Lösung vorrechnen/erklären können.

  Falls Sie die Lösung nicht ausreichend erklären können, werden Sie entweder ganz von der Abgabe der Hausaufgaben ausgeschlossen (d.h. nur noch Formel II ist anwendbar), oder es werden nur die Punkte dieses Blattes auf 0 gesetzt (Entscheidung der Übungsleiterin).
- Für die Übung herrscht Anwesenheitspflicht, damit das Vorrechnen tatsächlich geprüft werden kann.

  Sie dürfen maximal 3 Mal ohne Entschuldigung fehlen. Danach liegt es allein im Ermessen der Übungsleiterin, ob sie Ihre Leistung noch anerkennt (wenn Sie z.B. ansonsten sehr aktiv in der Übung waren).

Übungsschein (4)

Hausaufgaben, Fortsetzung:

- Falls die Anwesenheitspflicht für Sie unzumutbar ist, melden Sie sich bitte beim Dozenten.

  Sie sollten das Problem möglichst frühzeitig klären (bevor Sie drei Mal gefehlt haben). Für maximal fünf Teilnehmer kann eine alternative mündliche Prüfung angesetzt werden zum Nachweis, daß Sie die Aufgaben selbst gelöst haben.
- "Zu ähnliche" Lösungen verschiedener Gruppen haben ernste Konsequenzen (Täuschungsversuch).

  Punkteabzug bis zu 100%. Es ist gut, einander Tipps bei speziellen Problemen zu geben. Die fertige Lösung zum Abschreiben zu geben, nimmt die Chance zur eigenen Beschäftigung mit der Aufgabe.

Übungsschein (5)

Für Wirtschaftsinformatiker:

- Die erreichte Prozentzahl entspricht nicht genau den Fachpunkten.
- Garantiert bestanden (50 FP) haben Sie mit 60%.

  Genauer gesagt wird der Dozent nach Durchsicht der Klausur einen Wert u zwischen 48% und 60% festlegen, und einen Wert o zwischen 88% und 95%. Die untere Grenze u wird in 50 Fachpunkte umgerechnet, die obere Grenze o in 95 Fachpunkte. Falls Sie x Prozent erreicht haben, $u \leq x \leq o$, bekommen Sie $50 + 45 * (x-u)/(o-u)$ Fachpunkte. Bei $x < u$ bekommen Sie $50 * (x/u)$ Fachpunkte. Bei $x > o$: Individuelle Festlegung, bei $x \geq 98$: 100 FP. Garantierte Mindestzensuren also:

Prozent $\geq$	60	64	68	72	76	80	83	87	91	95
FP $\geq$	50	55	60	65	70	75	80	85	90	95
Zensur $\leq$	4.0	3.7	3.3	3.0	2.7	2.3	2.0	1.7	1.3	1.0

Übungsschein (6)

Beispiel:

- Ein Student hat in den Hausaufgaben insgesamt 80 von 100 möglichen Punkten erreicht (80%), und in der Klausur 20 von 40 möglichen Punkten (50%).
- Außerdem hat er einen Tippfehler im Skript gefunden und dafür einen Extrapunkt (1%) erhalten.
- Da die Hausaufgaben besser als die Klausur sind, wird Formel I angewendet: $80*0.3+50*0.7 = 59\%$.
- Durch den Extrapunkt wird das Ergebnis auf 60% erhöht, damit ist das Bestehen gesichert (50 FP).

Sollte der Dozent $u = 48$ und $o = 88$ wählen, ergeben sich 64 FP.

Übungsschein (7)

Nachholklausur:

- Voraussichtlich am 28.03.2007, 10^{00}–12^{00}.

  Bitte Aushänge wegen eventueller Verschiebungen beachten. Bei der Nachholklausur werden die Hausaufgabenpunkte auch berücksichtigt (falls besser als Klausur). Sie können die Hausaufgabenpunkte aber nicht in zukünftige Semester übertragen.

Für Bioinformatiker:

- Den Schein bekommen Sie wie oben erklärt.
- Die eigentliche Vordiplomsprüfung ist mündlich.

  Prüfungen werden voraussichtlich in der Zeit 26.–28.02.2007 abgenommen (bitte Aushang beachten). Sie können auch einen anderen Termin bekommen, schlimmstenfalls müssen Sie 1–2 Monate warten.

Übungsschein (8)

Für Informatiker:

- Sie können diese Vorlesung in Diplomprüfungen einbringen. Diese Prüfungen sind mündlich.
- Der Schein ist keine Vorbedingung, aber natürlich eine gute Übung (Vorbereitung).

Für Wirtschaftsmathematiker:

- Wenn Ihre Prüfungsordnung einen (Übungs-)Schein vorschreibt, müssen Sie an der Klausur teilnehmen.

  Wir vergeben keine Scheine allein aufgrund der Hausaufgaben.

Punkte-Datenbank im WWW

- Fehler beim Notieren der Punkte für Hausaufgaben und Klausuren kommen vor.
- Deswegen geben wir Ihnen die Möglichkeit, Ihren Punktestand im WWW zu kontrollieren:

  [http://www.informatik.uni-halle.de/~brass/db06/#PUNKTE]

- Sie müssen sich in die Datenbank eintragen (und ein Passwort vergeben) bevor die ersten Hausaufgabenpunkte eingetragen werden.

  Sonst kann sich jemand anders unter Ihrem Namen registrieren.

Ansprechpartner (1)

Dozent: Prof. Dr. Stefan Brass

- Email: `brass@informatik.uni-halle.de`
- Büro: Von-Seckendorff-Platz 1, Raum 313
- Telefon: 0345/55-24740
- Sprechstunde: Freitags, 12^{00}–13^{00}
- Frühere Unis: Braunschweig, Dortmund, Hannover, Hildesheim, Pittsburgh, Gießen, Clausthal.
- Oracle8 Certified Database Administrator.
- IBM Certified Database Associate (DB2 UDB 8.1)

Ansprechpartner (2)

Übungsleiterin: Dr. Annemarie Herrmann

- Büro: Von-Seckendorff-Platz 1, Raum 315
- Telefon: 0345/55-24737
- Email: `herrmann@informatik.uni-halle.de`

Sekretärin: Ramona Vahrenhold

- Büro: Von-Seckendorff-Platz 1, Raum 324
- Telefon: 0345/55-24750, Fax: 0345/55-27333
- Email: `vahrenhold@informatik.uni-halle.de`

Vorlesungs-Materialien

- ca. 1100 Folien.

  Die für die Klausur besonders relevanten Kapitel sind in Deutsch, eventuell einige der hinteren Kapitel noch in Englisch.

- Wird vor der Vorlesung auf die Webseite gestellt:
  `http://www.informatik.uni-halle.de/~brass/db06/`

  Formate: (1) PDF, farbig, gross, (2) Postscript, s/w, 4:1 verkleinert.

- Dort auch: Alte Klausuren, Links.
- Hausaufgaben auf Übungsseite:
  `http://www.informatik.uni-halle.de/dbs/lehre/DBI06_07/`
- Forum im StudIP: `http://studip.uni-halle.de`

Lehrbücher (1)

- Ramez Elmasri / Shamkant B. Navathe: Fundamentals of Database Systems, 4th Ed.

  Addison Wesley Longman, 2003, ISBN 0-8053-1755-4 (Hardcover), 1009 pages.

- Alfons Kemper / André Eickler: Datenbanksysteme. Eine Einführung. 5. Auflage.

  Oldenbourg, 2004, ISBN 3-486-27392-2, 640 Seiten.

- Raghu Ramakrishnan, Johannes Gehrke: Database Management Systems, 3rd Ed.

  McGraw Hill, June 2002, ISBN 0-07-246563-8, 1065 pages.

Lehrbücher (2)

- Joachim Biskup:
  Grundlagen von Informationssystemen.
  Vieweg, 1995, ISBN: 3-528-05494-8, 543 Seiten.
- Andreas Heuer, Gunter Saake:
  Datenbanken. Konzepte und Sprachen. 2. Aufl.
  MITP Bonn, 2000, ISBN 3-8266-0619-1, 704 Seiten.
- Gottfried Vossen:
  Datenbankmodelle, Datenbanksprachen und Datenbankmanagementsysteme. 4. Aufl.
  Oldenbourg, 2000, 3-486-25339-5, 778 Seiten. 5. Aufl. 2007!

Lehrbücher (3)

- Christopher J. Date:
  An Introduction to Database Systems, 8th Ed.
  Addison-Wesley, 2003, ISBN 0-321-18956-6.
- Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom:
  Database Systems: The Complete Book.
  Prentice Hall, 2001, ISBN 0-13-031995-3, 1152 pages.
- Philip M. Lewis, Arthur Bernstein, Michael Kifer:
  Database and Transaction Processing, 2nd Ed.
  Addison-Wesley, 2005, ISBN 0-321-26845-8, 1235 pages.

Lehrbücher (4)

- Chris J. Date, Hugh Darwen:
  A Guide to SQL Standard, 4th Edition.
  Addison-Wesley, April 1997, ISBN 0-201-96426-0, 530 pages.
- James R. Groff, Paul N. Weinberg:
  SQL: The Complete Reference, 2nd Ed.
  Osborne/McGraw-Hill, Sept. 2002, ISBN 0-07-222559-9, 1000 pages.
  CD-ROM (SQL Server, IBM DB2, MySQL) (Test Lizenzen).
- Jim Melton, Alan R. Simon: SQL: 1999 — Understanding Relational Language Components.
  Morgan Kaufmann, Mai 2001, ISBN 1558604561, 928 pages.

Software (Oracle)

- Verbreitetes DB-Management**system** (DBMS).
  Software zur Implementierung von anwendungsspezifischen DBen.
- Im Windows 2000 Pool, Raum 3.34.
  Sie müssen Windows/Unix-Login haben, Anträge neben Raum 3.35. Einen Oracle-Login bekommen Sie von uns, wenn Sie sich in die Punkte-DB eintragen. Die Datenbank heißt `studdb`.
- Alternativ: `ssh` nach `bach.informatik.uni-halle.de`.
- Oder auf eigenem PC (via OAI kostenlos).
  Oracle 9i läuft auf Windows NT/2000/XP Professional, Linux. Es benötigt 256 MB RAM, 2–3 GB Plattenplatz. Falls später Oracle Designer: Enterprise Version, sonst reicht Personal Version. Ältere Oracle-Versionen laufen auch auf anderen Windows-Versionen.

OAI Teilnahme (1)

- Öffnen Sie [http://oai.oracle.com].

  OAI steht für "Oracle Academic Initiative".

- Klicken Sie auf "Apply for Student Membership".
- Wählen Sie "Country=Germany", "Continue".
- Laden Sie den Vertrag (PDF Datei) herunter, und geben Sie ihn unterschrieben in der Übung ab.
- Klicken Sie auf "Weiter". Geben Sie bei "Einrichtung suchen" das Wort "Halle" ein, klicken Sie auf "Suchen", wählen Sie "MLU Halle-Wittenberg".

OAI Teilnahme (2)

- Wählen Sie die Abteilung "Institut für Informatik".
- Geben Sie Namen, EMail-Adresse und ein Passwort ein, und klicken Sie auf "Antrag absenden".

  Sie können sich anschliessend auf dem OAI-Server mit EMail-Adresse und Passwort einloggen. Es bietet sich an, auch dem Oracle Technology Network (OTN) beizutreten (kostet nichts, gibt Zugriff auf interessante Informationen).

- CD-Sätze zum Kopieren gibt es gegen ein Pfand von 3–5 Euro im Sekretariat (Raum 324).
  Oder Download von [http://technet.oracle.com].

Oracle Zertifizierung (1)

- Für Oracle 9i gibt es drei Stufen:
  - ◇ Oracle9i DB Administrator Certified Associate
  - ◇ Oracle9i DB Administrator Certified Professional
  - ◇ Oracle9i DB Administrator Certified Master
- Für erste Stufe muß man zwei Prüfungen bestehen, davon die erste über das Internet von zu Hause aus.

  Multiple Choice Tests. Die Prüfungen kosten jeweils ca. 100 Euro, Beispiel-Prüfungsfragen nochmal dasselbe. Rabatte durch OAI!
  Für zweite Stufe braucht man normalerweise einen Kurs von Oracle (ca. 1500 Euro), aber die Übungen zu DB I + DB II B zählen auch.

- Diese Vorlesung sollte Sie auf die erste Prüfung vorbereiten (keine Garantie).

Oracle Zertifizierung (2)

- Für Oracle 10g gibt es die gleichen drei Stufen.
- Allerdings gibt es nur noch eine Prüfung pro Stufe.

  Für die Associate Stufe ist das die Prüfung "Oracle Database 10g: Administration I": 79 Fragen in 105 Minuten, man braucht 68% zum Bestehen. Preis: 125$ (-40% OAI Rabatt).

- Die Vorlesung "Datenbanken II B" ist nützlich zur Vorbereitung auf die Zertifizierung, aber Sie müssen in jedem Fall noch selbständig weiter lernen.

  Z.B. die Handbücher lesen, die Beispielfragen von Self Test Software durchgehen, und das System selbst ausprobieren.

IBM DB2 Zertifizierung

- Auch drei Stufen:
  - ◇ IBM Certified Database Associate
  - ◇ IBM Certified Database Administrator
  - ◇ IBM Certified Advanced Database Administrator
- Eine Prüfung pro Stufe.

  Für die erste Stufe (Test 700 für DB2 UDB V8.1) z.B. 54 Fragen in 75 Minuten, man braucht 61% zum Bestehen. Preis für Studierende vermutlich 10 Euro.
- Für die erste Stufe ist ein großer Teil der Prüfung durch die Inhalte dieser Vorlesung abgedeckt.

  Es gibt eine Arbeitsgemeinschaft zur Vorbereitung auf die Prüfung.

Software (Andere DBMS)

- Die Hausaufgaben können nicht nur mit Oracle, sondern auch mit anderen DBMS gelöst werden.

  Es könnte eventuell leichte Probleme mit den Beispiel-Datenbanken geben. Das DBMS muß mindestens den vollen SQL-86 Sprachumfang unterstützen. Sie dürfen keine den Tutoren unbekannten systemspezifischen Erweiterungen des SQL-2003 Standards verwenden.
- Von welchen DBMS haben Sie schon gehört?
  - ◇ _____
  - ◇ _____
  - ◇ _____

Ausstattung der DB-Gruppe

- Sun Fire V440 (4 CPUs, 16 GB RAM)

  ca. 14.500 Euro
- Sun Fire V490 (4 Doppelkern-CPUs, 16 GB RAM)

  ca. 30.000 Euro
- Eine kleinere Sun, zwei Linux-Server
- SAN (Storage Area Network) mit
  - ◇ Fibre Channel RAID System: $36 \times 146\,\text{GB} = 5.2\,\text{TB}$

    ca. 36.000 Euro
  - ◇ SATA RAID System: 12×400 GB $= 4.8$ TB

    ca. 13.000 Euro

Verbesserung der Lehre

- Diese Vorlesung soll kein Monolog werden: Fragen sind sehr willkommen.
- Gute Lehre ist für mich wichtig. Ich möchte ein Datenbank-Lehrbuch schreiben.
- Vorschläge zur Verbesserung der Vorlesung sind sehr willkommen.
- Korrekturen für Fehler auf den Vorlesungs-Materialien, nützliche Links für die WWW-Seite etc. werden eventuell mit Extrapunkten belohnt.