Datenbanken I

(Winter 2008/2009)

Prof. Dr. Stefan Brass

Institut für Informatik

Übungsleiter: Dipl.-Inform. Christian Goldberg

Themen

- Grundlegende Begriffe, DB-Software (Funktionen)
- Etwas Logik, Konzepte des relationalen Modells

  Man kann viele Datenmodelle und Datenbanksprachen mit den Begriffen der Logik verstehen. Logik hilft, SQL besser zu nutzen.
- 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, XML.

Ansprechpartner (1)

Dozent: Prof. Dr. Stefan Brass

- Email: `brass@informatik.uni-halle.de`

  Betreff-Zeile sollte mit `[db08]` beginnen, möglichst aussagefähig.
- Büro: Von-Seckendorff-Platz 1, Raum 313
- Telefon: 0345/55-24740
- Sprechstunde: Dienstags, 12^{00}–13^{00}
- Frühere Unis: Braunschweig, Dortmund, Hannover, Hildesheim, Pittsburgh, Gießen, Clausthal.
- Oracle8 Certified Database Administrator.
- IBM Certified Advanced DBA (DB2 UDB 8.1).

Ansprechpartner (2)

Übungsleiter: Dipl.-Inform. Christian Goldberg

- Büro: Von-Seckendorff-Platz 1, Raum 315
- Telefon: 0345/55-24776
- Email: `goldberg@informatik.uni-halle.de`

Sekretärin: Ramona Vahrenhold

- Büro: VSP. 1, Raum 324 (dienstags geschlossen)
- Telefon: 0345/55-24750, Fax: 0345/55-27333
- Email: `vahrenhold@informatik.uni-halle.de`

Zeit und Ort (1)

Vorlesung (4 SWS):

- Montags, 10^{15}–11^{45}, Raum 3.28.
- Mittwochs, 8^{15}–9^{45}, Raum 3.28.

Anwesenheit nicht verpflichtend (Folien/Skript), aber:

- Eventuell wichtige Ankündigungen.
- Beispiele, Aufgaben, Diskussionen, Fragen.
- Zeitaufwand für selbst nacharbeiten eher höher.

  Wenn man nicht regelmäßig dranbleibt, wird man schnell "abgehängt".

Zeit und Ort (2)

Übung (2 SWS):

- Tafelübung, Hausaufgaben besprechen (~45 min), anschließend betreute Zeit am Rechner (~45 min).

  Genaue Zeitaufteilung je nach Thema variabel.

- Drei Gruppen (Anmeldung: StudIP):

ID	Tag	Zeit	Raum	Beginn
DI14	Dienstags	14^{15}–15^{45}	0.03/3.34	21.10.
DO12	Donnerstags	12^{15}–13^{45}	0.03/3.32	23.10.

Zeitliche Belastung

- Diese Vorlesung hat 10 Leistungspunkte.

  Auch "credit points" genannt.

- Entspricht 300 Stunden studentischer Arbeitszeit:

Lernform	SWS	Stunden
Vorlesung	4	60
Selbststudium	0	120
Tafelübung	1	15
Praktische Übung (betreut)	1	15
Praktische Übung (unbetreut)	(1)	15
Lösen von Hausaufgaben	0	35
Spezielle Prüfungsvorbereitung	0	40

Modulanmeldung (1)

- Für fast alle Studiengänge ist die Modulanmeldung über die Selbstbedienungsfunktion CSS Pflicht.

  Z.B. BA Informatik 180, BA Bioinformatik, BA Wirtschaftsinformatik 180. Falls nicht über CSS, dann im Prüfungsamt. [http://www.verwaltung.uni-halle.de/SBStud/]

- Die Modulanmeldung ist 14 Tage vor Vorlesungsbeginn bis 14 Tage nach Vorlesungsbeginn möglich, in diesem Semester also 22.09.2008 bis 20.10.2008.
- Die Modulanmeldung ist zwingende Voraussetzung für die spätere Anmeldung zur Modulleistung.

Modulanmeldung (2)

- Außerdem muß man sich bei StudIP anmelden.

  Mindestens um einen Platz in einer Übungsgruppe zu bekommen. Außerdem wird basierend auf der StudIP-Anmeldung eine Liste in der Vorlesung zur Modulanmeldung umlaufen (? redundant).

- Bei dieser Vorlesung wird drittens noch um die Anmeldung bei meiner Punkte-DB gebeten.

  [http://www.informatik.uni-halle.de/~brass/db08/#PUNKTE]
  Sie können so Ihre Hausaufgaben-Punkte einsehen, und später auch zur Klausur detaillierte Auswertungen finden. 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! Wenn Sie aus Datenschutz-Gründen nicht in der Datenbank stehen wollen, teilen Sie es mir mit.

Modulanmeldung (3)

Modulvoraussetzungen (Bedingung für Teilnahme):

- Nur wünschenswert, nicht obligatorisch.

  Ohne Voraussetzungen: Teilnahme auf eigene Gefahr.

- Programmierkenntnisse, z.B. erfolgreicher Abschluß der Vorlesung "Objektorientierte Programmierung".

  SQL ist eine sehr spezielle Programmiersprache und wird in dieser Vorlesung von Grund auf definiert. Konzepte von Programmiersprachen wären aber zum Verständnis nützlich. Im Kapitel über Anwendungsprogrammierung müßten Sie programmieren können. Konzepte wie Klassen wären auch zum Verständnis des ER-Modells nützlich.

- "Grundlagen und Konzepte der Modellierung" oder "Einführung in die Wirtschaftsinformatik".

Prüfung (1)

Klausur (geplant für 30.03.2009, 10^{00}–12^{00}):

- 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.

  Diesmal eventuell auch Aufgabe(n) mit kurzem Aufsatz.

- Nachklausur: Während des Sommersemesters.

  Der genaue Termin wird noch bekanntgegeben.

Prüfung (2)

Modulvorleistung (Bedingung für Klausurteilnahme):

- Mindestens 50% der Hausaufgabenpunkte.

  Es wäre unklug, nach 50% aufzuhören (Klausur-Vorbereitung!).

- Aktive Mitarbeit in den Übungen (s.u.).

Anmeldung zur Modulleistung (Klausur):

- Soweit ich weiß (ohne Garantie, informieren Sie sich bei dem für Sie zuständigen Prüfungsamt!), müssen Sie sich zur Klausur bis vier Wochen vorher anmelden (CSS/Prüfungsamt).

  Geht nur, wenn zum Modul angemeldet und Vorleistung erfüllt.

Prüfung (3)

Hausaufgaben:

- Bearbeitung in Gruppen von 1–4 Personen.

  Es können nur Teilnehmer der gleichen Übungsgruppe (gleicher Termin) Hausaufgaben zusammen abgeben. Von gemeinsam erarbeiteten Lösungen bitte nur ein Exemplar abgeben (nur eine EMail)!

- Ausgabe: Montags in Vorlesung/Netz,
  Abgabe: Montag der nächsten Woche.

  Abgabe elektronisch per Email an `goldberg@informatik.uni-halle.de` (bis 13^{15}, nur ASCII oder PDF, Betreff-Zeile muß "`DB08`", die Nummer des Übungsblattes, die ID der Übungsgruppe (z.B. `DI14`), und die Nachnamen aller Gruppenmitglieder enthalten. Bitte schicken Sie ASCII-Lösungen (z.B. für SQL-Anfragen) als Text der EMail (nicht als Anhang). Für ER-Entwurf ggf. Abgabe auf Papier.

Prüfung (4)

Hausaufgaben, Fortsetzung:

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

  Falls Sie die Lösung nicht ausreichend erklären können, können Sie ganz von diesem Modul ausgeschlossen werden (wegen Täuschungsversuch). Eventuell werden auch nur die Punkte dieses Blattes auf 0 gesetzt (Entscheidung des Übungsleiters).

- 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 des Übungsleiters, ob sie Ihre Leistung noch anerkennt (wenn Sie z.B. ansonsten sehr aktiv in der Übung waren).

Prüfung (5)

Hausaufgaben, Fortsetzung:

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

- Sie können Gruppen während des Semesters wechseln oder bei einer Hausaufgabe einmal aussetzen (wenn Sie an der Lösung nicht beteiligt waren).

- "Zu ähnliche" Lösungen verschiedener Gruppen haben ernste Konsequenzen (Täuschungsversuch).

  Z.B. keine Punkte für alle beteiligten Gruppen. Es ist gut, einander Tipps bei speziellen Problemen zu geben. Die fertige Lösung zum Abschreiben zu geben, nimmt die Chance zum eigenen Lernen.

Prüfung (6)

Zur Benotung:

- Falls die Noten nach Ihrer Prüfungsordnung in Zahlen von 0 bis 100 ausgedrückt werden, ist nicht garantiert, daß die erreichte Prozentzahl der Klausurpunkte genau der Note entspricht.

- Es ist möglich, daß für besondere Aktivitäten Extrapunkte vergeben werden (bis maximal 10% der Gesamtpunktzahl der Klausur).

  Nach Ermessen des Dozenten und soweit mit der Prüfungsordnung verträglich. Z.B. wichtige Hinweise auf Fehler auf den Folien, oder andere gut begründete Verbesserungsvorschläge, auffallend aktive Mitarbeit in Vorlesung und/oder Übung, etc. Typischerweise nur 1–3%.

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/db08/`

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

- Dort auch: Alte Klausuren, Links.

  Sowie Folien des vorigen Semesters. Werden eventuell aktualisiert.

- Forum im StudIP: `http://studip.uni-halle.de`

Lehrbücher (1)

- Alfons Kemper / André Eickler:
  Datenbanksysteme. Eine Einführung. 6. Auflage.

  Oldenbourg, 2006, ISBN 3-486-57690-9, 672 Seiten, 39.80 Euro.

- Andreas Heuer, Gunter Saake:
  Datenbanken. Konzepte und Sprachen. 2. Aufl.

  MITP Bonn, 2000, ISBN 3-8266-0619-1, 704 Seiten.
  Dritte Auflage geplant für Okt. 2007, mit Kai-Uwe Sattler, 39.95.

- Gottfried Vossen:
  Datenbankmodelle, Datenbanksprachen und Datenbankmanagementsysteme. 4. Aufl.

  Oldenbourg, 2000, 3-486-25339-5, 778 Seiten. 5. Aufl. Okt. 2007!

Lehrbücher (2)

- Ramez Elmasri / Shamkant B. Navathe:
  Fundamentals of Database Systems, 5th Ed.

  Addison Wesley, 2006, ISBN 032141506X (Softcover, 78.06),
  ISBN 0321369572 (Hardcover, 104.81 Euro), 1168 pages.
  Deutsche Ausgabe Grundstudium (Teile fehlen/nur online): Pearson Studium, 3. Aufl., 2005, ISBN 3827371538, 550 Seiten, 29.95 Euro.

- Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom:
  Database Systems: The Complete Book.

  Prentice Hall, 2001, ISBN 0-13-031995-3, 1152 pages, 87.98 Euro.
  Taschenbuch: 2007, ISBN 0-13-098043-9, 73.78 Euro.
  Neuauflage angekündigt für 2009 (?).

Lehrbücher (3)

- Christopher J. Date:
  An Introduction to Database Systems, 8th Ed.

  Addison-Wesley, 2003, ISBN 0-321-18956-6, 1024 Seiten, 83.41 Euro.

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

  McGraw Hill, June 2002, ISBN 0-07-115110-9 (1065 Seiten, 67.75 Euro), 0-07-246563-8 (1104 Seiten, 142.13 Euro).

- Philip M. Lewis, Arthur Bernstein, Michael Kifer:
  Database and Transaction Processing, 2nd Ed.

  Addison-Wesley, 2005, ISBN 0-321-26845-8, 1272 pp., 88.35 Euro.
  TB: ISBN 0-321-31256-2, 76.99 Euro. Gibt "Introductory Version".

Lehrbücher (4)

- Chris J. Date, Hugh Darwen:
  A Guide to SQL Standard, 4th Edition.

  Addison-Wesley, 1997, ISBN 0-201-96426-0, 544 pages, 62.01 Euro.

- James R. Groff, Paul N. Weinberg:
  SQL: The Complete Reference, 2nd Ed.

  Osborne/McGraw-Hill, Sept. 2002, ISBN 0-07-222559-9, 1080 pages. CD-ROM (SQL Server, IBM DB2, MySQL) (Test Lizenzen), 57.92.

- Jim Melton, Alan R. Simon: SQL: 1999 — Understanding Relational Language Components.

  Morgan Kaufmann, 2. Aufl., Mai 2001, ISBN 1558604561, 928 pages, 63.23 Euro.

Software (Oracle)

- Verbreitetes DB-Managementsystem (DBMS).
- Im PC Pool (3.32), Thin Client Pool (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 10g läuft auf Windows 2000/XP Professional/Server 2003, Linux (SUSE Enterprise, Professional mit etwas Arbeit). Es benötigt 512 MB RAM, 5 GB Plattenplatz. Falls später Oracle Designer: Enterprise Edition, sonst beliebige Edition. Oracle 10g Express Edition: auch ohne OAI kostenlos (sogar für echte Nutzung).
  [http://www.oracle.com/technology/products/database/xe/]
  Linux-Tipps: [http://www.puschitz.com/OracleOnLinux.shtml]

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

- Für Oracle 10g gibt es drei Stufen:
  - ◇ Oracle10g DB Administrator Certified Associate

    Multiple-Choice-Test "Oracle Database 10g: Administration I": 79 Fragen in 105 Minuten man braucht 68% zum Bestehen. Preis: 125$ (-40% OAI Rabatt).
  - ◇ Oracle10g DB Admin. Certified Professional

    Für diese Stufe braucht man normalerweise einen Kurs von Oracle (ca. 1500 Euro), aber die Übungen zu DB I + DB IIB zählen vermutlich auch.
  - ◇ Oracle10g DB Administrator Certified Master
- Die Vorlesungen DB I und DB IIB sind Hilfe bei der Vorbereitung, aber eigenes Lernen ist nötig. Ggf. Arbeitsgemeinschaft.

IBM DB2

- Anderes wichtiges, sehr verbreitetes DBMS.
- Kann für die Übungen zu dieser Vorlesung genauso wie Oracle verwendet werden.
- Die " Express-C Edition" ist kostenlos: "Free to develop, deploy, distribute".

  Läuft auf Windows und Linux (32 bit und 64 bit). Einschränkungen bei der Nutzung größerer Hardware (mehr als 2 CPU Kerne, 2 GB RAM), keine Garantie, kein Support (es gibt aber Forum/Cummunity), keine Updates, Verbot von Clustern.
- IBM DB2 wird in Kürze auch in den Pools zur Verfügung stehen.

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 730 für DB2 UDB V9) z.B. 64 Fragen in 90 Minuten, man braucht 59% zum Bestehen. Preis für Studierende vermutlich 10 Euro. Zuletzt ganz von IBM gesponsort.
- Für die erste Stufe ist ein großer Teil der Prüfung durch die Inhalte dieser Vorlesung abgedeckt.

  Aber dennoch ist ca. eine Woche selbständiges Lernen nötig.

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: $36 \times 146\,\mathrm{GB} = 5.2\,\mathrm{TB}$

    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.

  Vorschläge zur Verbesserung der Vorlesung sind sehr willkommen!

- Ich möchte dieses Semester in der Vorlesung mehr Aufgaben, Gruppenarbeit etc. machen.

  Falls sich nicht mindestens ein Drittel der Teilnehmer melden, rufe ich auch Teilnehmer auf, die sich nicht melden.

- Bei manchen Teilen werde ich Sie bitten, daß Sie sich den Stoff anhand Folien selbst erarbeiten.

  Wir werden dann nur noch über Fragen sprechen.

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)
  - ◇ ______
  - ◇ ______
  - ◇ ______
  - ◇ ______
  - ◇ ______