INFSCI 2710: Database Management

Summer 2000

University of Pittsburgh, School of Information Sciences

Stefan Brass

Main Topics of This Course

1. Data Models (Entity-Relationship and Relational) and Data Modeling (Introduction to Database Design).
2. Relational Database Management Systems like Oracle, Solid Knowledge of the Database Language SQL.
3. Database Programming (Overview).
4. Data Protection (Integrity Constraints, Backup), Transactions, Concurrency Control.
5. Outlook, e.g. Deductive Databases (time permitting).

Things You Know Already (1)

UNIX, Solaris, Linux	☐ not at all	☐ a bit	☐ well
C	☐ not at all	☐ a bit	☐ well
Pascal	☐ not at all	☐ a bit	☐ well
C++	☐ not at all	☐ a bit	☐ well
Java	☐ not at all	☐ a bit	☐ well
SQL	☐ not at all	☐ a bit	☐ well
Microsoft Access	☐ not at all	☐ a bit	☐ well
Oracle	☐ not at all	☐ a bit	☐ well
Other DBMS:	____________________		

Things You Know Already (2)

B-Trees	☐ not at all	☐ a bit	☐ well
Multitasking, Locks	☐ not at all	☐ a bit	☐ well
Context-Free Grammars	☐ not at all	☐ a bit	☐ well
Syntax Graphs	☐ not at all	☐ a bit	☐ well
Regular Expressions	☐ not at all	☐ a bit	☐ well
Sets, $\cup$, $\cap$, $\times$	☐ not at all	☐ a bit	☐ well
Logic, $\wedge$ (and), etc.	☐ not at all	☐ a bit	☐ well

About This Course

- Lectures: Wednesdays, 6:00 – 8:50 PM, 169 Crawford Hall
- Instructor: Stefan Brass
  Phone: (412) 624-9404
  Email: `sbclass@sis.pitt.edu` (or `sbrass@...`)
  Room: SIS 726
  Office Hours: Mondays, 5:00 – 7:00 PM
- GSA: Ms. Siripun Sanguansintukul
  Email: siripun@lis.pitt.edu
  Office Hours: Wednesdays, 4:00-5:45 PM (B203, IS Bldg.)

Textbooks For This Course

- Ramez Elmasri, Shamkant B. Navathe:
  Fundamentals of Database Systems, Third Edition.
  Addison-Wesley, 1999, ISBN 0-8053-1755-4,
  960 pages, ca. $65–107 (Hardcover).

Alternatives:

- Ramez Elmasri, Shamkant B. Navathe:
  Fundamentals of Database Systems, Second Edition.
- Avi Silberschatz, Henry Korth, S. Sudarshan:
  Database System Concepts, Third Edition.

Optional:

- R. Sunderraman: Oracle Programming.
  Addison-Wesley, 2000, ISBN 0-201-61258-5, 304 pages.

Using Database Systems: Oracle I

- SIS Solaris Systems, Room 819.
- Telnet access: SQL*Plus has no graphical interface.
  You can use it via telnet to `paradox.sis.pitt.edu`
  However, reading the documentation might be difficult.
- You can get Oracle Software through the Software Licensing Services of the University, 203 Bellefield Hall ($5 per CD).
  See `http://www.technology.pitt.edu/software/`.
  Personal Oracle for Windows NT/95/98 includes the Oracle Server, SQL*Plus, SQL Navigator, ODBC Support, Documentation in HTML Format, etc.
- See also `http://oraclestore.oracle.com/`.
  30-day trial versions: $39.95 per package. Download free.

Using Database Systems: Oracle II

- The newest version is currently (May 2000) Oracle 8i, Release 2 (Oracle 8.1.6).
  But this is at the moment only available as enterprise version for NT. And the Software Licensing Department does not give enterprise versions to students.
- Personal Oracle 8i (8.1.5) is available for Windows NT and for Windows 98 (and for Linux from the Oracle Website).
  The personal version is a full RDBMS, the restrictions only refer to parallel processing, partioning, object extensions, etc.
- The Oracle Client CD contains a precompiler for using SQL in C and the Enterprise Manager (a GUI-Tool for the DBA).
- Which Oracle 8 version you use is not important for this course.

Using Database Systems: Other DBMS

Other Commercial Systems:

- 60-day trial version of IBM DB2 Universal Database for free: `http://www.software.ibm.com/data/db2/udb/`
- Microsoft Access: SIS Pentium Lab, Room 838.

Systems Available with Source Code, Free for Academia:

- Mini-SQL: `http://www.Hughes.com.au/`
  Runs e.g. on Linux, only subset of SQL.
- MySQL: `http://mysql.he.net/`
- PostgreSQL: `http://www.postgresql.org/`
  Runs e.g. on Linux, large SQL subset + OO extensions.

Database Vendors (1)

- Market for databases in 1998: 7.1 Billion \$ (+15%)

Vendor	Market Share	Change to 1997
IBM	32.3%	+3.4%
Oracle	29.3%	−0.1%
Microsoft	10.2%	+0.3%
Informix	4.4%	−0.4%
Sybase	3.5%	−1.0%
Others	20.5%	−2.2%

Source: Dataquest. In Computer-Zeitung No 13, Apr. 1, 1999.

- Performance Benchmark Results: `http://www.tpc.org/`

Database Vendors (2)

Database Market Share (1998) [Dataquest, March 1999]

Vendor	RDBMS	UNIX	NT
Oracle	38.5%	60.9%	46.1%
IBM	30.8%	7.3%	10%
Microsoft	7%	—	29.7%
Informix	6%	13%	—
Sybase	5%	7%	3%
Others	14%	11.8%	12%
Market Size:	\$5.4 Billion	\$2.2 Billion	\$1.2 Billion
Growth Rate:	+18%	+10%	+46%

See `http://www.oracle.com/database/marketshare/index.html?/database/marketshare/sld002.htm`

Information in the World Wide Web

`http://www2.sis.pitt.edu/~sbrass/db00_2/`:

- Updated table of contents, all slides and other materials.
- Information about:
  - Free and commerical database software.
  - Selected database books.
  - Some influential database researchers.
  - Some database conferences.

See also: `http://www.informatik.uni-trier.de/~ley/db/`
`http://www.acm.org/sigmod/`

About Grading (1)

Homeworks:

- There will be homeworks more or less every week.
- Homeworks are due in the next week later before the class.
  You may submit up to three homeworks up to one week late.
  However, no submission will be accepted after April 20, 6pm.
- Homeworks are an important preparation for the exams!
- Homeworks can be solved in groups of 1–3 persons.
- Use email submissions only if you cannot attend a lecture.
  Send them directly to `siripun@lis.pitt.edu`
- Questions about homeworks: `sbclass@sis.pitt.edu`
  Resend if no answer in 24h. Please check the web page first.

About Grading (2)

Homework Results in the Web:

- You can see your current homework results in the web:
  `http://www2.sis.pitt.edu/~sbrass/grades/view.html`
- Please register in
  `http://www2.sis.pitt.edu/~sbrass/grades/reg.html`
- Choose a password which you don't use in any other place.
- You must register, even if you do not intend to use this service. Otherwise, somebody else can register for you.
- Web databases are not very secure. If you do not want your results in the database, contact me. We cannot be made liable for security holes.

About Grading (3)

Distribution of Points:

- 35% Homework Assignments
- 25% Midterm Exam (February 23)
  Exams will be "open book" (but without using computers)
- 40% Final Exam (April 26)
  The better exam is weighted 40%, the worse one 25%.
- 2% extra: SQL puzzle announced during the lecture.
  I reserve the right to give a few students which have been especially active during the class or have suggested important improvements for the transparencies (e.g. error corrections), up to 5% of extra points.

About Grading (4)

Grading Scale:

- ≥99.0%: A+
- 90–98%: A
- 85–89%: A−
- 80–84%: B+
- 75–79%: B
- I reserve the right to lower these boundaries slightly.
  There might be an opportunity to do an extra project in order to improve the grade (maybe also after the course is finished via a "G" grade). But this must be limited to a very small number of students and will be my own decision.

About Grading (5)

Curving:

- No curve will be applied. If all course participants have 90% of the points, all will get an A grade.
- However, I guarantee that
  - at least the best 20% will get an A grade,
  - the next 15% will get an A− (or better),
  - and the next 20% will get a B+ (or better).
- Previous results have been much better, but the students also have worked very hard.

Certification Programs

- Oracle 8 Certified Database Administrator:
  You have to pass 5 exams. This course should cover about 90% of the material of the first exam (SQL, PL/SQL).

  See `http://education.oracle.com/certification/`. You can download a free assessment test. INFSCI 2711 gives further preparation. However, major self-study is still necessary. You must work with Oracle at home (as DBA) and study the Oracle documentation. You should also read “Couchman: Oracle 8 Certified Professional. Osborne, 1999”.
- IBM DB2 Certification:
  `http://www.software.ibm.com/data/db2/udb/db2cert.html`

  See also: C. Janacek, D. Snow: DB2 Universal Database V5 Certification Guide, Prentice Hall, ISBN 0-13-079661-1.

Exercise

Name some typical applications of databases, e.g. companies or organizations which use them, or any private uses you might have for a database:

1. ______________________________
2. ______________________________
3. ______________________________
4. ______________________________
5. ______________________________

Acknowledgements

I would like to thank:

- My academic teachers, especially:
  Prof. Dr. Udo Lipeck and Prof. Dr. H.-D. Ehrich.
- All the students from my courses. Especially the following students told me about errors or suggested improvements:

  Prisda Gomutputra, Melanie K. Huber, Samuel Jackendoff, Steve Krelis, Jong-Hoon Lee, Cuiwei Li, Shelly Lukon, Jerry Maher, Lisa Meng, Stella Noh, Theresa Pappas, Susan Poole, Fei Qin, Nicholas Steadman, Hikmet Umar, Julie Zheng Zhou, Bing Zhu.
- My wife Nina, who really supported me and gave me all the time I needed to work on the transparencies.