

Oberseminar "Datenbanken" WS 2002/03

Nachfolgend finden Sie einige Hinweise und Fragestellungen zu den ausgegebenen Themen. Die mit * gekennzeichneten Materialien sind leihweise bei mir bzw. auch als Kopie erhältlich.

Die angegebenen Oracle Links im Internet (und weitere) können Sie über meine Homepage -> Datenbank-Links erreichen.

Datenbanken für Entscheidungsunterstützung

1. Data Mining Grundlagen (2 Vorträge mit Ausarbeitung)

Data Mining macht es möglich, interessante Pattern in Unternehmensdatenbeständen oder E-Commerce-Daten herauszufinden. Diese Pattern liefern oft wertvolle Erkenntnisse für Marketingzwecke oder zur Schwachstellenanalyse. Eine häufige Anwendung ist dabei Customer Relationship Management.

Die Vorträge sollten folgende Schwerpunkte umfassen:

- Motivation und Anwendungen
- Eingaben: Konzepte, Instanzen, Attribute
- Ausgabe: Wissensdarstellung
- Basis-Algorithmen des Data Mining (Lernalgorithmen): Ableitung von Regeln, Entscheidungsbäume, Clustering u.a.
- Weitere Probleme des Data Mining, z.B. Aufbereitung der Ein- und Ausgabe

Quellen:

- I. Witten, E. Frank: Data Mining - Praktische Werkzeuge und Techniken für das maschinelle Lernen, Hanser-Verlag, 2001.*
- R. Ramakrishnan, J. Gehrke: Database Management Systems, 2nd Edition, McGrawHill, 1999 (Chapter 24: Data Mining)

2. Oracle Data Mining (1 Vortrag - mit Implementierung)

Die Oracle Data Mining Suite ist schwerpunktmäßig ein Werkzeug für Customer Relationship Management. Es erlaubt die Segmentierung von Kunden und Voraussagen über ihr Verhalten. Dies dient u.a. der Gewinnung neuer Kunden, besseres Eingehen auf vorhanden Kundenbedürfnisse (Profiles), Cross-Selling, Erkennung von Betrugsfällen. Es läßt sich leicht mit vorhandenen Datenbanken kombinieren und gestattet die parallele Ausführung von Data Mining Algorithmen, Import und Export von Daten ist sehr leicht möglich. Die Mining-Algorithmen sind: Classification und Regression Trees (CART), Neuronale Netzwerke, k-nearest neighbor, Regression und Clustering. Der Vortrag sollte auf folgende Aspekte eingehen:

- Verwendete Basis-Algorithmen im Oracle Data Mining
- Prototypische Implementierung einer Daten-Analyse (mit selbstgewähltem Beispiel) mit Erfahrungsbericht

Quellen:

- R. Ramakrishnan, J. Gehrke: Database Management Systems, 2nd Edition, McGrawHill, 1999 (Chapter 24: Data Mining)
- L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone: Classification and Regression Trees, Wadsworth, Belmont, 1984.
- CD-ROMs Oracle Data Mining * (wahlweise auch für Linux erhältlich)
- Oracle Technology Network: <http://www.oracle.com/products/bi>
-

3. Aufbau von Data Warehouses (1 Vortrag mit Ausarbeitung)

Beim Aufbau von Data Warehouses sind eine Reihe konzeptioneller Probleme zu lösen, die sich aus den Anforderungen an die Performance, Datenqualität und Aktualität der Daten im Warehouse ergeben. Bei der Wahl des Beispiels ist eine Abstimmung mit dem nachfolgenden Vortrag über OLAP erforderlich. Der Vortrag sollte auf folgende Aspekte eingehen:

- Aufbau eines Data Warehouses
- Entwurfsphasen bei einem Data Warehouse
- Probleme der Datenintegration (Data Cleansing)
- Konsistenzaspekte
- Neue SQL-Operationen zur Unterstützung des ETL-Prozesses in Oracle9i
- Produkte in Oracle (Oracle Warehouse Builder)

Quellen:

- Bauer, Günzel: Data Warehouse Systeme - Architektur, Entwicklung, Anwendung, dpunkt-Verlag, 2001.*
- W. Inmon u.a.: Managing the Data Warehouse, John Wiley & Sons, 1996.
- R. Ramakrishnan, J. Gehrke: Database Management Systems, 2nd Edition, McGrawHill, 1999 (Chapter 23: Decision Support)*
- Oracle-Kurs Data Warehousing *
- Kurs Oracle9i SQL und PL/SQL (ptm-Akademie)*
- Oracle Technology Network: <http://www.oracle.com/products/bi>

4. Online Analytical Processing/OLAP (1 Vortrag mit Ausarbeitung)

OLAP basiert auf einer multidimensionalen Datenbank und führt Analysen und Berechnungen durch, um What-If-Szenarien und andere Entscheidungsunterstützung zu ermöglichen. Der Vortrag sollte folgende Aspekte behandeln:

- wichtigste OLAP-Operationen und deren Anwendung
- OLAP-Unterstützung in Oracle-SQL
- Behandlung von Materialized Views
- Produkte in Oracle (Oracle Discoverer)

Quellen:

- Bauer, Günzel: Data Warehouse Systeme - Architektur, Entwicklung, Anwendung, dpunkt-Verlag, 2001.*
- Kähler: SQL mit Oracle, Vieweg Verlag (Kapitel 11) od. andere Quelle über SQL materialized Views und OLAP Operatoren
- R. Ramakrishnan, J. Gehrke: Database Management Systems, 2nd Edition, McGrawHill, 1999 (Chapter 23: Decision Support)*
- W. Inmon: Managing the Data Warehouse, John Wiley & Sons, 1996.
- Oracle Technology Network: <http://www.oracle.com/products/bi>

Non-Standard-Datenbanken - Neue Datentypen

5. Objektorientale Datenbanken (1 Vortrag mit Ausarbeitung)

Der SQL3 Standard beschreibt die Erweiterungsmöglichkeiten von relationalen Datenbanken um objektorientierte Konzepte. Diese Konzepte sind vorzustellen und mit den Möglichkeiten von Oracle zu vergleichen. Schwerpunkte sind:

- Large Objects
- User Defined Types (UDTs)
- Typ-Konstrukturen: strukturiertere und Collection-Typen
- User Defined Methods, Functions und Procedures
- Typed Tables und Views

Quellen:

- U. Hohenstein, V. Pleßer: Oracle 8 Effiziente Anwendungsentwicklung mit objektorientierten Konzepten, dpunkt Verlag, 1998.*
- A. Christiansen, M. Höding, C. Rautenstrauch, G. Saake: Oracle 8 effizient einsetzen, Addison Wesley, 1998.*
- Online Dokumentation Oracle 8i Spatial Reference and User's Guide (bzw. Oracle
- SQL3 Tutorial* oder SQL-Norm und Java*

6. Text-Datenbanken: Oracle InterMedia Text (1 Vortrag mit Implementierung)

Ein großer Teil der Daten liegen in semistrukturierter (z.B. HTML) oder unstrukturierter Form (z.B. Text) vor. Aus diesem Grund ergibt sich die steigende Bedeutung von Volltext-Datenbanken. Während in der Vergangenheit eigenständige Systemlösungen existierten, ist nun ein

Zusammenwachsen mit klassischen Datenbanksystemen zu beobachten. Am Beispiel von Oracle InterMedia Text sollen die Möglichkeiten von Volltext-Datenbanken demonstriert werden:

Der Vortrag sollte auf folgende Aspekte eingehen:

- Laden von Texten in Datenbanken
- Indexieren und Index-Wartung
- Anfragen und Anfrageoperatoren, insb. "interessante" Text Retrieval Operationen
- Highlighting und Präsentation von Ergebnisdokumenten
- Vergleich Oracle InterMedia Text mit SQL/MM Full Text

Quellen:

- Online Dokumentation Oracle 8i interMedia Text Reference
- SQL3 Tutorial (SQL/MM Text) *
- <http://www.oracle.com> (White Papers u.a.)
- Oracle Kurs interMedia Text *
- Oracle Technology Network: <http://technet.oracle.com/products/text>

7. Multimediale Datenbanken mit Oracle interMedia Audio, Image, Video (1 Vortrag zu Audio, 1 Vortrag zu Image, jeweils mit Implementierung)

Oracle8i interMedia Audio, Image and Video ist eine Komponente von Oracle8i interMedia. Sie erlaubt die Behandlung von Bildern, Tönen und Videos in einer integrierten Weise mit anderen Unternehmensinformationen über Objekt-Typen (benutzerdefinierte Datentypen). DBMS werden angewandt auf multimediale Inhalte im Inthnet, E-Commerce und multimedia-haltige Anwendungen.

Die Fähigkeiten von interMedia Audio, Image, Video umfassen Speicherung, Retrieval und Management multimedialer Daten in Oracle8i. Dazu zählen im einzelnen:

InterMedia ist so entworfen, daß es erweiterbar ist. Es unterstützt eine Basis-Menge von gebräuchlichen Audio-, Image- und Video-Daten-Eigenschaften, die erweitert werden können, z.B. zur Unterstützung zusätzlicher Formate oder neuer Kompressions-Algorithmen.

InterMedia ist ein Baustein für verschiedene Multimedia-Applikationen und keine Endbenutzer-Anwendung. Es besteht im wesentlichen aus Objekt-Typen zusammen mit entsprechenden Methoden zur Verwaltung und Verarbeitung multimedialer Daten, z.B.:

- Internet-Musikläden, die Samplings in CD Qualität anbieten
- Digitale Bildergalerien
- Verwaltung elektronischer Dokumente (Document Imaging)
- Digitale Video-Bibliotheken und -Archive

Diese Anwendungen haben gemeinsame und spezifische Eigenschaften. Oracle8i interMedia Audio, Image, Video Objekt-Typen unterstützen allgemeine Anforderungen und können für spezifische Anforderungen erweitert werden. Oracle8i interMedia ist in JAVA, C++, oder traditionellen Programmiersprachen.

InterMedia unterstützt die Speicherung vieler bekannter Dateiformate. Es liefert Mittel, um Audio-, Image- und Video-Spalten oder Objekte zu existierenden Tabellen hinzuzufügen, multimediale Daten zu speichern und zu lesen und liefert begrenzte Konvertierungsmöglichkeiten zwischen Anwendungsformaten. Dies erlaubt Datenbank-Designern, existierende DB-Anwendungen um multimediale Daten zu erweitern oder neue Multimedia-DB-Anwendungen zu bauen. Der Vortrag sollte auf folgende Aspekte eingehen:

- Multimediale Datentypen: Audio, Video, Image
- Operationen auf multimedialen Daten
- Inhaltsbasierte Suche bei Image-Daten
- Entwicklungsumgebung Media Developers Kit (MDK)

Quellen:

- Online Dokumentation: Oracle 8i interMedia Audio, Image, Video Reference and User's Guide
- Online Dokumentation: Oracle8i Visual Information Retrieval User's Guide and Reference
- Oracle Technology Network: <http://technet.oracle.com/products/intermedia>

8. Räumliche Datenbanken: Oracle Spatial (1 Vortrag mit Ausarbeitung)

Diese Technologie erlaubt neue Arten von Enterprise-Applikationen, vor allem Geographische Informationssysteme (GIS), die in vielen Anwendungsbereichen wie Telekommunikation, Verkehr, Administration, Einzelhandel und Versorgung benötigt werden. Das in Oracle Spatial zugrundeliegende Modell folgt dem allgemeinen Open GIS-Standard, Oracle Spatial kooperiert mit anderen Werkzeugen zur Verarbeitung räumlicher Daten (z.B. Visualisierungssoftware).

Räumliche Datenbanken wie Oracle Spatial Data erlauben die Speicherung und die Abfrage räumlicher Daten. Der Vortrag sollte auf folgende Aspekte eingehen:

- Räumliche Datentypen: z.B. Punkt, Polygon
- Zugriffsstrukturen für Punkte und Regionen, z.B. Grid Files, R-Trees
- Operationen auf räumlichen Daten
- Vergleich Oracle Spatial mit SQL/MM Spatial
- Implementation eines selbstgewählten Anwendungsbeispiels in Oracle Spatial

Quellen:

- Online Dokumentation Oracle 8i Spatial Reference and User's Guide
- SQL3 Tutorial (SQL/MM, Object-Relational) *
- W. Kim (Ed.): Advanced Database Systems *
- Oracle Technology Network: <http://technet.oracle.com/products/spatial>

9. Temporale Datenbanken mit Oracle Time Series (1 Vortrag)

Oracle8i Time Series (vormals Time Series Cartridge) ist eine Erweiterung von Oracle 8i, die die Speicherung und Retrieval von zeitbehafteten Daten mit Hilfe von Objekttypen zur Verfügung stellt. Time Series ist ein Baustein für die Anwendungsentwicklung und somit keine Endbenutzer-Anwendung. Es besteht aus Datentypen mit zugehörigen Funktionen zur Verarbeitung von Zeitreihendaten.

Zum Beispiel können Anwendungen Time Series verwenden, um historische Daten, die von Finanzmarkt-Transaktionen abgeleitet sind (z.B. Börsengeschäfte, Kurse), zu verarbeiten. In solchen Anwendungen können mit Hilfe von Time Series-Funktionen bestimmte Operationen bequem durchgeführt werden:

- Ermittlung von Aktienkursen an einem bestimmten Datum
- Berechnung monatlicher Umsätze für eine Aktie in einem bestimmten Jahr
- Berechnung gleitender 30-Tage-Durchschnitte während eines Jahres

Time Series-Anwendungen haben besondere Anforderungen und ein bestimmtes Maß an Gemeinsamkeiten. Diese Gemeinsamkeiten sind in den Time Series-Datentypen ausgedrückt, die aber auch Erweiterbarkeit unterstützen. Es können bestehende Tabellen genutzt bzw. angepaßt bzw. auch neue Tabellen angelegt werden.

Der Vortrag sollte auf folgende Aspekte eingehen:

- Basiskonzepte temporaler Datenbanken: Gültigkeitszeit/Aufzeichnungszeit, temporale Datentypen, Kalender
- Konsistenz zeitbehafteter Datenbanken
- Architektur von TimeSeries
- Vergleich TimeSeries mit Temporal SQL / Standardisierungsbemühungen SQL-3

Quellen:

- Online Dokumentation: Oracle8i Time Series User's Guide
- W. Kim (Ed.): Advanced Database Systems, Chapter Temporal Database Systems *
- Kristian Torp, Christian S. Jensen, Michael Böhlen: Layered Implementation of Temporal DBMSs - Concepts and Techniques, 1997, TIMECENTER Technical Report TR-2
- Richard T. Snodgrass, Michael Böhlen, Christian S. Jensen, Andreas Steiner: Transitioning Temporal Support in TSQL2 to SQL3, 1997, TIMECENTER Technical Report TR-8
- TimeCenter: www.timecenter.com
- TimeConsult: www.timeconsult.com

XML und Datenbanken

10.XML-Unterstützung in Oracle 9i/Release 2 (1 Vortrag mit Ausarbeitung)

Es ist ein Vortrag zu halten über die XML-Unterstützung von Oracle, wobei die neuen Möglichkeiten von O9i Release 2 zu behandeln sind:

- XML Datentyp in 9i R2 für "native" XML-Objekte

- Alternativen zur Speicherung von XML in Oracle 9i
- Anfragen auf XML-Objekten
- Beispiel-Anwendung

Quellen:

- <http://www.oracle.com/xml>
- <http://www.doag.de/sig/tools>: Vortrag *Relational and beyond, die Oracle XML Datenbank*

11.Speicherung von XML-Dokumenten im Oracle Internet File System (IFS) (1 Vortrag mit Implementierung)

Das Internet File System (iFS) erlaubt die Speicherung und den Zugriff auf Dokumente auf Grundlage einer file-basierten Metapher. Es ist Bestandteil von Oracle 8.1.7 (oder höher). Das iFS erleichtert die Arbeit mit XML und kann die folgenden Aufgabe für XML-Dokumente erledigen:

- Automatisches Parsen von XML und Abspeicherung des Inhalts
- Darstellen (Rendern) der XML-Dateiinhalte zur Anzeige oder Verarbeitung auf der Client-Seite

Die Vorträge sollten folgende Schwerpunkte beinhalten:

- iFS Architektur
- Funktionsweise aus Datenbanksicht
- Anwendungen, insb. XML Framework (Parser, Renderer)

Quelle:

- Oracle Technology Network: <http://technet.oracle.com/products/ifs>

Verteilte und Mobile Datenbanken

12.Mobile Datenbanken (1 Vortrag mit Ausarbeitung)

Oracle unterstützt mit seinem Produkt Oracle 9i Lite die Nutzung mobiler Datenbanken auf Laptops oder Organizern. Es ist ein Überblicksvortrag zu halten, der die grundlegenden Konzepte sowie die Entwicklung mobiler DB-Anwendungen behandelt.

- Replikation von Daten / Synchronisationskonzepte
- Administration: System Management, User und Device Management
- Connectivity: Protokolle
- Zuverlässigkeit / Skalierbarkeit

Quelle:

- Oracle Technology Network: <http://technet.oracle.com/products/lite>

13.Oracle Advanced Queueing (1 Vortrag mit Implementierung)

Oracle Advanced Queueing (AQ) unterstützt den asynchronen Austausch von Daten zwischen Anwendungen über Messages, die in Queues abgelegt werden. Anhand einer Beispiel-Anwendung ist der Gebrauch von AQ zu demonstrieren, z.B. beim Versenden von XML-Nachrichten, Integration von E-Mail.

- Grundkonzepte der asynchronen Kommunikation
- AQ-Package von Oracle
- Demonstration des Beispiels

Quelle:

- Oracle Technology Network: <http://technet.oracle.com/products/aq>

Java und Datenbanken

14.Java und Datenbanken (1 Vortrag mit Ausarbeitung)

Java-Anwendungen benötigen eine Datenbank-Anbindung, für die JDBC ein Standard ist. Es ist ein Überblick zu geben über die Java-Unterstützung von Oracle 9i sowie die Weiterentwicklung des JDBC-Standards.

- Java Database Connectivity (JDBC): Architektur und APIs
- Überblick JDBC 1.2 und 2.0, Ausblick auf 3.0
- Java-Unterstützung in Oracle (JVM, Stored Procedures)

Quellen:

- S. DeBloch: SQL-Norm und Java, Datenbank Tutorientage 2001.*
- W. Dehnhardt: Anwendungsprogrammierung mit JDBC, Hanser-Verlag, 1999.*
- G. Saake, U Sattler: Datenbanken und Java. JDBC, SQLJ und ODMG, dpunkt-Verlag 2000.
- Oracle Technology Network: <http://technet.oracle.com/java>

15.Oracle TopLink und Java Data Objects (1 Vortrag mit Ausarbeitung)

Ein objektorientiertes Programmiermodell benötigt transparente Objekt-Persistenz, die auf der Basis von JDBC die Anbindung beliebiger relationaler Datenbanken unterstützt. Es sind die konzeptionellen Probleme bei der objekt-relationalen Abbildung darzustellen sowie gegenwärtige Stand bei der Entwicklung von Standards und Produkte.

- Abbildungsproblem Java Klassen / Relationen
- Architektur und Vergleich von Oracle TopLink und JDO

Quellen:

- Java Data Objects (JDO), in Datenbank-Spektrum 3/2002*
- Oracle Technology Network: <http://technet.oracle.com/products/ias/toplink>