

Oberseminar "Datenbanken" WS 2001/02

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 mit Oracle (3 Vorträge)

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. Aufbau von Data Warehouses (1 Vortrag)

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 Vortrag über OLAP erforderlich. Der Vortrag sollte auf folgende Aspekte eingehen:

- Aufbau eines Data Warehouses
- Metadaten, Common Warehouse Model (CWM)

- Entwurfsphasen bei einem Data Warehouse
- Probleme der Datenintegration (Data Cleansing)
- Konsistenzaspekte / Materialized Views

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)*
- <http://www.oracle.com/> (=>Warehouse)
- Oracle-Kurs Data Warehousing *

3. Online Analytical Processing (OLAP) (1 Vortrag)

OLAP basiert auf einer multidimensionalen Datenbank und führt Analysen und Berechnungen durch, um What-If-Szenarien und andere Entscheidungsunterstützung zu ermöglichen. Es sind die wesentlichen Verfahren und deren Anwendungen vorzustellen. Als Beispiel-System kann Oracle EXPRESS vorgestellt werden.

Quellen:

- Bauer, Günzel: Data Warehouse Systeme - Architektur, Entwicklung, Anwendung, dpunkt-Verlag, 2001.*
- 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.
- <http://www.oracle.com/> (=>Warehouse)

Non-Standard-Datenbanken - Neue Datentypen

4. Objektorientale Datenbanken (1 Vortrag)

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: strukturierte und Collection-Typen
- User Defined Methods, Functions und Procedures
- Typed Tables und Views

Quellen:

- U. Hohenstein, V. Pleßer: Oracle 8 Effiziente Anwendungsentwicklung mit objektrelationalen 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
- SQL3 Tutorial* oder SQL-Norm und Java*

5. Text-Datenbanken: Oracle InterMedia Text (1 Vortrag)

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
- 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 *

6. Räumliche Datenbanken: Oracle Spatial (1 Vortrag)

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 *
- <http://www.oracle.com> (White Papers u.a.)

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

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 Inhternet, 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

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
- <http://www.oracle.com>

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

Oracle8i Time Series (vormals Time Series Cartridge) ist eine Erweiterung von Oracle 8i, das 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 mit 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

Die Vorträge unter der Rubrik "XML und Datenbanken" sind von den Vortragenden inhaltlich abzustimmen (z.B. bezüglich gemeinsamer Grundlagen). Praktische Beispiele zur Veranschaulichung der Konzepte sind wünschenswert. In einigen Vorträgen sind die von Oracle verfügbaren Werkzeuge vorzustellen.

9. XML-Datenbankstandards (1 Vortrag)

Es ist ein Überblick über die zur Zeit vorhandenen Anfragesprachen und die aktuellen Standardisierungsbemühungen zu geben. Folgende Schwerpunkte sind hierbei zu setzen:

- Anfragesprachen: XPath, XML-QL, XQL, XQuery (XQuery)
- DTD, XML Schema

10. Speicherung von XML-Dokumenten in (objekt)relationalen Datenbanken (2 Vorträge)

Die Speicherung von XML-Dokumenten als Daten empfiehlt sich, wenn das XML-Dokument eine wohldefinierte Struktur hat und Daten enthält, die geändert oder auf andere Weise genutzt werden können. Typischerweise weisen solche XML-Dokumente eine komplexe Struktur auf (z.B. Kaufaufträge, Flugpläne). Es soll ein Überblick über vorhandene Algorithmen gegeben werden, wie XML-Dokumente auf relationale Datenbanken gemappt werden können.

Im folgenden sollen speziell die objektrelationalen Erweiterungen von Oracle 8i betrachtet werden. Oracle bietet die Möglichkeit, die Struktur der Daten in der Datenbank mittels Objekttypen, Objektreferenzen und Collections zu erfassen. Es gibt zwei Optionen zur Speicherung und Bewahrung der Struktur in einer objektrelationalen Form.

- Speicherung der Element-Attribute in einer relationalen Tabelle und Definition von Objekt-Views, um die Struktur der XML-Elemente zu erfassen
- Speicherung der strukturierten XML-Elemente in einer Objekt-Tabelle

Die Vorträge sollten folgende Schwerpunkte beinhalten:

- Mapping-Algorithmen XML-relational: inhaltsorientiert und strukturorientiert
- objektorientierte Ansätze: Object Views, objektrelationale Speicherung

11. Speicherung von XML-Dokumenten als Large Objects (opaker Ansatz) (1 Vortrag)

Die Speicherung von XML-Dokumenten als Dokument in einem Large Object ist brauchbar bei statischem Inhalt (z.B. Artikel, Bücher, Verträge), der sich nur bei einem Ersetzen des ganzen Dokuments ändert. Somit nutzt man die Stärken von Datenbanksystemen bei der Speicherverwaltung aus. Selbst bei Speicherung außerhalb einer Oracle-Datenbank (z.B. BFILE) kann man die Möglichkeiten der Text-Indexierung nutzen. Der Vortrag sollte auf folgende Aspekte eingehen:

- Einsatz von Oracle interMedia Text für XML
- Oracle XML Developer Kit (XDK) für Large Objects
- Arten von Anfragen: interMedia Text vs. XPath

12. Entwicklung von XML-Anwendungen mit Oracle (1 Vortrag)

In diesem Vortrag soll die dynamische Erstellung von XML-Dokumenten aus Oracle-Datenbanken mit Hilfe des XML SQL Utility erläutert werden, ebenso die Eingabe von Daten über XML.

13. Dynamische Publikation von Datenbankinhalten mittels XSQL Pages (1 Vortrag)

In diesem Vortrag sollen die Möglichkeiten von Oracle XSQL Pages als Publishing Framework demonstriert werden, auch im Zusammenwirken mit anderen Komponenten zur flexiblen Formatierung der Ergebnisse.

Quellen (für die Themen 9-13)

- S. Muench: Building Oracle XML Applications, O'Reilly, 2000.*
- H. Schöning: XML und Datenbanken, Tutorium, 2001.*
- Ron Bouret: XML & Databases; <http://www.rpbouret.com/xml/XMLAndDatabases.htm>
- T. Krumbein: Speicher- und Anfragemethoden für XML-Dokumente ohne Schema in objektrelationalen Systemen am Beispiel Oracle, Diplomarbeit, HTWK Leipzig, 2001 (enthält Verweise auf zahlreiche Publikationen)*
- M. Conrad: Diplomarbeit XML-Datenbanken, Dez.01 (für Thema 10)
- J. Timm: Speicherung von XML-Dateien in Objekt-relationalen Datenbanken, Diplomarbeit, Uni Rostock, 1999.*

14. Das XML-Datenbanksystem Tamino (1 Vortrag)

Der Vortrag soll die technische Architektur und die Grundkonzepte des nativen XML-Datenbanksystems Tamino der Software-AG vorstellen.

Quellen

- www.softwareag.com/tamino: White Papers u.a.

Java und Datenbanken

15. Java und Datenbanken (1 Vortrag)

Java-Anwendungen benötigen eine Datenbank-Anwendung, für JDBC ein Standard ist, auf dessen Basis vielfältige Client-Server-Architekturen entwickelt werden können.

- Java Database Connectivity (JDBC): Architektur und APIs
- Überblick JDBC 1.2 und 2.0, Ausblick auf 3.0
- SQLJ

Quellen:

- S. Deßloch: 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.

16. Enterprise Java Beans und Oracle 8i

Enterprise JavaBeans (EJB) ist eine Architektur zur Entwicklung transaktionaler Anwendungen als verteilte Komponenten in Java. EJB ist zugleich eine mächtige Entwicklungsmethodik für verteilte Anwendungsentwicklung. Enterprise Beans ersparen es dem Entwickler, sich um bestimmte Details zu kümmern, wie Transaktionssupport, Security, Remote Object Access und andere fehleranfällige Dinge. Dies wird dem Entwickler transparent durch den EJB-Server und einen Container zur Verfügung gestellt. EJB-Applikationen werden vollständig in Java entwickelt und benötigen keine zusätzliche Sprache. Session Beans realisieren applikatorische Logik.

Oracle8i JServer implementiert die EJB Version 1.0-Spezifikation. EJB spezifiziert Java Remote Method Invocation (RMI) als Transportprotokoll. Oracle8i JServer implementiert RMI über IIOP, das CORBA Internet Inter-ORB Protokoll. Oracle 8i gestattet damit direkten objekt-orientierten Zugriff auf eine steigende Anzahl von offenen Systemen. Der Vortrag sollte auf folgende Aspekte eingehen:

- Kurze Einführung in Enterprise Java Beans: Konzepte und Begriffe, Architektur, Programmierung
- Objektzugriff (JNDI), Verbindungsmanagement (IIOP), Security/Zugriffskontrolle
- Transaktionsmanagement
- Werkzeuge
- EJB-Beispielprogramm (evtl. Prototypimplementierung mit JServer)

Quellen:

- Online Dokumentation: Oracle8i Enterprise JavaBeans and CORBA Developer's Guide
- <http://www.oracle.com>; technet

-