

Bericht von der GI-Fachtagung BTW '97 – Datenbanken in Büro, Technik und Wissenschaft, Ulm, 5.–7. März 1997

S. Conrad*

W. Hasselbring†

Die 1985 in Karlsruhe begonnene, im zweijährlichen Turnus veranstaltete Tagungsreihe “BTW” hat es sich zum Anliegen gemacht, Datenbank-Forscher, Hersteller und Anwender zusammenzuführen, um gemeinsam den Entwicklungsstand und die Perspektiven neuer Datenbank-Technologien in aktuellen Einsatzgebieten zu diskutieren. Die BTW fand in diesem Jahr an der Universität Ulm statt und wurde von knapp 200 Teilnehmern besucht.

Insgesamt wurden 82 Beiträge eingereicht. Von diesen wurden 16 Langbeiträge und 7 Kurzbeiträge akzeptiert, so daß die Akzeptanzrate für Langbeiträge bei ca. 20% lag. Wie Klaus Dittrich (Vorsitzender des Programmkomitees) in der Begrüßung betonte, handelte es sich dabei, insbesondere auch im Vergleich zu vielen internationalen Konferenzen, um eine relativ niedrige Akzeptanzrate, die auf eine hohe Qualität der akzeptierten Beiträge schließen läßt.

1 Eingeladene Vorträge

M. Kersten (Uni Amsterdam/CWI) bot in seinem Vortrag über *Research and Business Challenges in Data Mining Technology* einen schönen Überblick über die Forschung im Bereich *Data Mining* und *Data Warehouse* sowie seinen *Data Surveyer* Ansatz.

A. Meier (CSS Versicherung Luzern/ETH Zürich) berichtete in seinem Vortrag *Der Einsatz objektorientierter Datenbanksysteme in der Praxis: ein Triathlon* über das Schwimmen (Analyse existierender Systeme), das Radfahren (Aufbau eines Entwicklungsprozesses und einer Entwicklungsumgebung) und das Laufen (den Betrieb) beim Einsatz der objektorientierten Datenbank O₂ in einem Projekt bei der CSS Versicherung Luzern. Das Hauptproblem war zunächst die Performance, das Projekt war jedoch ein Erfolg. Wichtig war auch die Verbindung zur R3-basierten Altanwendung.

Ph.A. Bernstein (Microsoft) berichtete über die Arbeiten bei Microsoft zu *Repositories and Object Oriented Databases*.

2 Technische Beiträge

Da einige Vorträge parallel zu anderen Veranstaltungen liefen, möchten wir hier nur auf einige technische Beiträge kurz eingehen.

Das aktuelle Forschungsgebiet *Data Warehouses* und *OLAP* war nicht nur Thema des eingeladenen Vortrags von M. Kersten, sondern auch Gegenstand der folgenden Beiträge:

- M.-C. Wu und A. Buchmann: Research Issues in Data Warehousing
- P. Baumann, P. Furtado, R. Ritsch, N. Widmann: Anfrageformulierung und Ablage dimensionsbehafteter Daten in RasDaMan
- G. Höfling, M. Blaschka, B. Dinter, P. Spiegel und T. Ringel: Data Warehouse-Technologie für das Diagnosedaten-Management
- W. Lehner, M. Teschke und H. Wedekind: Aufbau und Auswertung multidimensionaler Daten

Somit kann dieses Thema als Schwerpunkt der Tagung identifiziert werden. Zusammen mit dem eingeladenen Vortrag von M. Kersten zeigen diese Vorträge, daß die Datenbankforschung sich den aktuellen Herausforderungen stellt, die als Problemstellungen gegenwärtig in der Wirtschaft eine große Rolle spielen.

J. Frohn, R. Himmröder, P. Kandzia, G. Lausen und C. Schlepphorst stellten FLORID vor, ein Prototyp-System für die objektorientierte, deduktive Datenbanksprache F-Logik (Frame-Logik). Aufgrund

*Universität Magdeburg, Fakultät für Informatik, Institut für Technische Informationssysteme, Postfach 4120, D-39016 Magdeburg, conrad@iti.cs.uni-magdeburg.de

†Universität Dortmund, Informatik 10, Software-Technologie, D-44221 Dortmund, willi@ls10.informatik.uni-dortmund.de

der sprachlichen Mächtigkeit der F-Logik stellt die Implementierung eines solchen Systems eine große Herausforderung dar. Zur Demonstration der Fähigkeiten dieser Logik, insbesondere der Möglichkeit, Daten und Metadaten gleichermaßen zu behandeln, wurde ein Beispiel aus dem Bereich der Schemaintegration gewählt. Aufgrund der getroffenen Vereinfachungen und Annahmen muß die Wahl des Beispiels allerdings als nicht überzeugend und damit als unglücklich bezeichnet werden.

In einem stark theoretisch geprägten Beitrag stellten A. Scharnofske, U. Lipeck und M. Gertz eine Erweiterung der „klassischen“ graphischen Anfragesprache QBE (Query By Example) um Unteranfragen sowie Aggregations- und Mengenoperationen vor. Gerade in Hinblick auf die Entwicklung immer leistungsfähigerer graphischer Anfrageoberflächen könnte diese Arbeit richtungsweisenden Charakter haben, da sie eine adäquate und intuitive Anfrageformulierung gerade für nicht fachlich vorgebildete Datenbank-Benutzer ermöglicht.

In dem Beitrag von S. Berchtold, D. Keim und H.-P. Kriegel wurde ein Verfahren zur Ähnlichkeitssuche in CAD-Datenbanken erläutert. Dieses Verfahren (*Section Coding* genannt) bestimmt für 2-dimensionale graphische Objekte (Polygone) eine größennormierte translationsunabhängige und annäherungsweise rotationsunabhängige Darstellung, die für einen effizienten Vergleich bei Ähnlichkeitsanfragen genutzt werden kann. Die Anzahl der Sektoren, in die ein Objekt zerlegt wird, bestimmt einerseits den Aufwand der Berechnung dieser Darstellung und andererseits den Fehler gegenüber der idealen Rotationsunabhängigkeit. Dieser Fehler ist dann auch maßgeblich für die Zuverlässigkeit bei der Beantwortung von Ähnlichkeitsanfragen.

W. Wilkes stellte in einem Kurzbeitrag das Rollen-Beziehungs-Modell vor, das generische Beziehungen in objektorientierten Datenbanken erlaubt. Allerdings wurde nicht klar, inwieweit sich dieser Vorschlag von anderen unterscheidet, die schon früher ein Rollenkonzept neben der üblichen Spezialisierungsbeziehung in objektorientierten Modellen vorgeschlagen haben.

Die Ergebnisse einer umfangreichen Performanz-Studie an objektorientierten Datenbanksystemen wurden von U. Hohenstein, V. Pleßer und R. Heller präsentiert. Anhand spezieller Beispielanwendungen wurden drei objektorientierte Datenbanksysteme miteinander verglichen. Bei der Untersuchung wurde auch die Auswirkung der verschiedenen Tuning-Maßnahmen, die auf verschiedenen Ebenen ansetzen können (z.B. durch Anlegen von Indexen), beobachtet.

G. Saake, A. Christiansen, S. Conrad, M. Höding, I. Schmitt und C. Türker gaben in einem Kurzbeitrag einen Überblick über ein vom Land Sachsen-Anhalt gefördertes Forschungsprojekt zur Föderation heterogener Datenbanksysteme. Neben Schemaintegration und globaler Integritätssicherung ist die Einbeziehung von Daten aus dateibasierten Datenverwaltungssystemen ein wesentlicher Schwerpunkt. Hierbei steht die Gewinnung der Schemainformation aus einer Menge von Dateien im Vordergrund.

L. Brüchert, J. Zimmermann und A. Buchmann haben in ihrem Beitrag die verschiedenen Möglichkeiten des Einsatzes von Triggern in einem Wertpapier-Archivierungssystem untersucht. Durch vergleichende Performanz-Messungen sind sie zu einer Bewertung gelangt, die zeigt, daß einige Formen der Verwendung von Triggern in Datenbanken keine Performanz-Steigerung bewirken.

S. Conrad, I. Schmitt und C. Türker präsentierten einen formalen Ansatz, der es erlaubt, Integritätsbedingungen bei der Schemaintegration von den Ausgangsschemata auf das integrierte Schema zu übertragen. Bisher vorgeschlagene Schemaintegrationsansätze berücksichtigen explizite Integritätsbedingungen überhaupt nicht oder nur in einer sehr eingeschränkten Form. Der hier vorgestellte Ansatz erscheint geeignet, in andere Schemaintegrationsansätze integriert zu werden.

J. Reinert und H.P. Steiert berichteten über ihre Erfahrungen mit der Kombination aus ereignisgesteuerten Prozeßketten für die Analyse, Booch für den Entwurf und Smalltalk für die Implementierung einer datenintensiven Anwendung. Hier wurden insbesondere die Datenmodellierungsfähigkeiten der Booch-Notation kritisiert. Es stellt sich allerdings die Frage, ob die Wahl der Booch-Notation für diesen Kontext sinnvoll war.

Weitere technische Beiträge, auf die wir hier nicht näher eingehen, seien der Vollständigkeit halber aufgelistet:

- T. Rölleke und N. Fuhr stellten einen Ansatz zum probabilistisches Schließen in Datenbanken mit vagen Daten vor.
- G. Braunreuther, V. Linnemann und H.-G. Lipinski: Unterstützung von Computersimulationen durch objektorientierte Datenbanksysteme am Beispiel einer Anwendung aus der Medizin.
- U. Nink und N. Ritter: Database Application Programming with Versioned Complex Objects.
- J. Friebe: eine GeoServer-Architektur zur Bereitstellung geographischer Basisdaten im Internet.
- T. Zurek: Parallel Temporal Joins.

- N. Krivokapić: Synchronisation in verteilten Objektsystemen.
- J. Sellentin und B. Mitschang: Möglichkeiten und Grenzen des Einsatzes von CORBA in DB-basierten Client/Server-Anwendungssystemen.
- W. Thoben: Sicherheitsanforderungen im Rahmen der Bedrohungs- und Risikoanalyse von IT-Systemen.
- J. Baader und M. Philipp: rechtliche Grundlagen für den Einsatz betrieblicher elektronischer Archivierungssysteme.

Aus dem Übersichtsvortrag von H. Loeser zur Datenbankanbindung an das WWW sei das folgende Zitat für die Nachwelt festgehalten: "Was früher der Fuchsschwanz am Manta war, ist heute die eigene Homepage im WWW".

3 Industrie-Sitzungen

Um Praxis und Forschung ins Gespräch zu bringen, wurden auf der BTW '97 erstmalig spezielle "Industrie-Sitzungen" durchgeführt, bei denen Hersteller und Anwender zusammen mit Wissenschaftlern über ihre Produkte bzw. Anwendungserfahrungen berichten und diskutieren. Es gab Produktvorstellungen für Data Warehouse, für objektrelationale und objektorientierte Datenbanksysteme und eine Sitzung zum Thema kleine und mittlere Unternehmen im WWW.

Diese Sitzungen wurden in Form von Podiumsdiskussionen durchgeführt. Der Moderator führte jeweils kurz in das Thema und seine verschiedenen Facetten ein, bevor die Firmen in Kurzpräsentationen (ca. 10 Min.) ihre Produkte und Tätigkeitsschwerpunkte darstellen konnten. Anschließend wurde mit den Firmenvertretern unter Einbezug des Publikums lebhaft diskutiert.

Die Einführung von speziellen Industrie-Sitzungen, die thematisch fokussiert sind, darf als Erfolg bewertet werden. Ob sie sich auf Dauer bewähren und den gewünschten Effekt haben, daß nach Zeiten rückläufiger Industriebeteiligung diese wieder merklich zunimmt, wird sich vielleicht schon auf der nächsten BTW herausstellen.

4 Weitere Veranstaltungen am gleichen Ort

In Verbindung mit der BTW'97 fand an der Universität Ulm am 4. März ein GI-Workshop zu „Multidimensionalen Datenbanken statt“. Schwerpunktthemen dieses Workshops waren Modellierung, Anfragesprachen (inkl. Anfrageoptimierung), Transaktionsmechanismen, Speicherverwaltung (inkl. Tertiärspeichereinbindung), Kompressionstechniken, Systeme (kommerzielle Produkte, Forschungsprototypen) sowie Erfahrungen aus Anwendungsgebieten. Veranstalter dieses Workshops war P. Baumann (FORWISS München).

Ebenfalls am 4. März fand ein Workshop über das DFG-Projekt „CROQUE (Cost and Rule based Optimization of object-oriented QUeries)“ statt. CROQUE ist ein von der DFG seit Anfang 1995 gefördertes gemeinsames Forschungsprojekt der Universitäten Konstanz und Rostock (Projektleiter: M. H. Scholl und A. Heuer). Im Mittelpunkt der Arbeiten stehen Untersuchungen zu Optimierungsverfahren und Verfahren des Query-Rewriting für regelbasierte und algebraische Anfragesprachen, Auswertungsverfahren für objektorientierte Anfragen sowie die Optimierung des physischen Entwurfs von Objekt-Datenbanken bezüglich eines gegebenen Anwendungsprofils. Im Rahmen des Workshops wurden aktuelle Ergebnisse dieses Projektes öffentlich vorgestellt.

5 Organisation

Die Tagung wurde durch das Organisationskomitee unter der Leitung von Peter Dadam kompetent und ohne Hektik durchgeführt. Besonders positiv hervorzuheben ist die Tatsache, daß der Teilnehmerschein zur Tagung die kostenfreie Benutzung der Busse und Straßenbahnen der Ulmer Nahverkehrsbetriebe ermöglichte. WWW-Information findet sich unter <http://hermes.informatik.uni-ulm.de/abt/dbis/BTW97/btw97.html>

6 Ausblick

Die nächste BTW soll voraussichtlich Anfang März 1999 in Freiburg im Breisgau stattfinden. Leiter der BTW'99 wird dann Georg Lausen sein.